

THE STOCHASTIC ENCOUNTER-MATING MODEL

ONUR GÜN AND ATILLA YILMAZ

ABSTRACT. We propose a new model of permanent monogamous pair formation in zoological populations comprised of $k \geq 2$ types of females and males, which unifies and generalizes the encounter-mating models of Gimelfarb (1988). In our model, animals randomly encounter members of the opposite sex at their so-called firing times to form temporary pairs which then become permanent if mating happens. Given the distributions of the firing times and the mating preferences upon encounter, which depend on the sex and the type of the animals, we analyze the contingency table $Q(t)$ of permanent pair types at any time $t \geq 0$. First, we consider definite mating upon encounter and provide a formula for the distribution of $Q(t)$. In particular, at the terminal time T , the so-called mating pattern $Q(T)$ has a multiple hypergeometric distribution. This implies panmixia which means that female and male types are uncorrelated in the expected mating pattern. Next, when the firing times come from Poisson and Bernoulli point processes, we formulate conditions that characterize panmixia. Moreover, when these conditions are satisfied, the underlying parameters of the model can be changed to yield definite mating upon encounter, and our results for the latter case carry over. Finally, when $k = 2$, we fully characterize heterogamy/panmixia/homogamy, i.e., negative/zero/positive correlation of same type females and males in the expected mating pattern. We thereby rigorously prove, strengthen and generalize Gimelfarb's results.

1. INTRODUCTION

1.1. The SEM model. Consider a zoological species comprised of $k \geq 2$ types of females and males forming permanent monogamous heterosexual pairs. Label the types $1, \dots, k$. Take a population consisting of $x_i \geq 0$ type- i females and $y_j \geq 0$ type- j males for every $i, j \in \{1, \dots, k\}$, such that

$$(1.1) \quad n := x_1 + \dots + x_k = y_1 + \dots + y_k \geq 1,$$

i.e., there are an equal number of females and males in the population. Denote the set of females (resp. males) by $Z_f = \{\varphi_1, \dots, \varphi_n\}$ (resp. $Z_m = \{\mu_1, \dots, \mu_n\}$), and the whole population by $Z := Z_f \cup Z_m$.

Associated to each $\zeta \in Z$, there is a simple point process $N(\zeta)$ on the real half-line $[0, \infty)$ with time points $(\tau_s(\zeta))_{s \geq 1}$. These point processes are mutually independent and the distribution of $N(\zeta)$ depends on the sex and the type of ζ . For the sake of convenience, we will refer to $(\tau_s(\zeta))_{s \geq 1}$ as the firing times of ζ . At these times, ζ initiates an encounter (i.e., forms a temporary pair) with a random member of the opposite sex. Whether such an encounter results in mating (i.e., the formation of a permanent pair) is also random, and the conditional probability of this event (upon encounter) depends on the types of the female and the male involved. Because of these two stages of encounter and mating, we will call this model of permanent pair formation *stochastic encounter-mating* (SEM).

Date: August 22, 2014; revised on May 11, 2015.

2010 Mathematics Subject Classification. 92D25, 60J28, 60G55.

Key words and phrases. Population dynamics, pair formation, encounter-mating, assortive mating, random mating, panmixia, homogamy, heterogamy, monogamy, mating preferences, mating pattern, contingency table, multiple hypergeometric distribution, simple point processes, Poisson process, Bernoulli process.

Let us now give a precise description of the model. We start with imposing certain conditions on the firing times.

DEFINITION 1.1. *A two-dimensional array $\mathbf{t} := (t_s(\zeta))_{s \geq 1, \zeta \in Z}$ of positive extended real numbers is said to be a proper family of firing times if it satisfies the following conditions:*

- (a) $t_s(\zeta) \leq t_{s+1}(\zeta)$ and the inequality is strict whenever $t_s(\zeta) < \infty$; and
- (b) $t_1(\zeta) < t_2(\zeta) < \dots < \infty$ for every $\zeta \in Z_f$ or for every $\zeta \in Z_m$.

The set of all proper families of firing times is denoted by Φ .

Our technical assumptions (see Subsection 2.1) on the distributions of the firing times will ensure that $\tau := (\tau_s(\zeta))_{s \geq 1, \zeta \in Z} \in \Phi$ almost surely.

For every $t \geq 0$, let $S_f(t)$ (resp. $S_m(t)$) denote the set of single females (resp. males) at that time. We will refer to $S(t) := S_f(t) \cup S_m(t)$ as the singles' pool. Initially, all of the animals are single, i.e., $S(0) = Z$. Given any $\mathbf{t} \in \Phi$, the distinct times at which at least one single animal fires are recursively defined as

$$(1.2) \quad t_r^* = t_r^*(\mathbf{t}) := \min\{t_s(\zeta) : s \geq 1, \zeta \in S(t_{r-1}^*), t_s(\zeta) > t_{r-1}^*\}$$

for $r \geq 1$, where $t_0^* := 0$ as a convention. These are the times of the so-called firing rounds. The encounter and mating stages at the r th firing round are as follows.

Stage I: Encounter. If exactly one element of $S(t_{r-1}^*)$, say a female, fires at t_r^* , then it samples a male from $S_m(t_{r-1}^*)$ uniformly at random and initiates an encounter with it. If two or more elements of $S(t_{r-1}^*)$ fire at t_r^* , then each of them samples a single member of the opposite sex uniformly at random and without replacement. Here is exactly how this is done: The elements of $S(t_{r-1}^*)$ that fire at t_r^* are ordered in an arbitrary way. (For the sake of definiteness, order them w.r.t. their labels and let the females go first if any. However, as we will see in Remark 2.2, this order does not matter.) Assume without loss of generality (WLOG) that the first one is a female $\varphi_a \in S_f(t_{r-1}^*)$ and it samples a male $\mu_b \in S_m(t_{r-1}^*)$. Then, neither φ_a nor μ_b can be sampled by the subsequent elements of $S(t_{r-1}^*)$ that fire at t_r^* . Moreover, even if μ_b fires at t_r^* , it is not allowed to sample a female from $S_f(t_{r-1}^*)$ when its turn comes, because it is already in a temporary pair with φ_a . This procedure continues until all the elements of $S(t_{r-1}^*)$ that fire at t_r^* are in a temporary pair. The collection of these pairs is denoted by $(\Delta\mathcal{L})'(t_r^*)$. This choice of notation will become clear in the next two paragraphs.

Stage II: Mating. After the encounter stage at the r th firing round is completed, independent Bernoulli random variables are assigned to the pairs in $(\Delta\mathcal{L})'(t_r^*)$ to determine whether they become permanent (i.e., if the courtship results in mating). The probability of mating for each such type- ij pair (i.e., if the female and the male are of type- i and type- j , respectively) is equal to some $p_{ij} > 0$. The $k \times k$ matrix

$$P := (p_{ij})$$

is called the *mating preference matrix* of the species. The collection of permanent pairs formed at the r th firing round is denoted by $\Delta\mathcal{L}(t_r^*) \subset (\Delta\mathcal{L})'(t_r^*)$. The singles' pool $S(t_r^*)$ is obtained by removing from $S(t_{r-1}^*)$ the animals in the pairs constituting $\Delta\mathcal{L}(t_r^*)$. Note that the animals in the pairs in $(\Delta\mathcal{L})'(t_r^*) \setminus \Delta\mathcal{L}(t_r^*)$, i.e., the ones that have formed temporary but not permanent pairs at the r th firing round, remain in the singles' pool.

This two-stage procedure is iteratively carried out at t_1^*, t_2^*, \dots and it naturally stops at

$$T := \min\{t > 0 : S(t) = \emptyset\} = \min\{t_r^* : r \geq 1, S(t_r^*) = \emptyset\},$$

i.e., when the singles' pool is depleted. Since $p_{ij} > 0$ and every female or every male fires infinitely many times in $[0, \infty)$ by Definition 1.1, the terminal time T is almost surely finite. For any $t \in [0, T]$, let

$$\mathcal{L}(t) := \bigcup_{\substack{r \geq 1: \\ t_r^* \leq t}} \Delta\mathcal{L}(t_r^*)$$

be the collection (or unordered list) of permanent pairs formed by time t . Similarly, for any $i, j \in \{1, \dots, k\}$, let $Q_{ij}(t)$ be the number of permanent type- ij pairs formed by time t . The

set-valued process $\mathcal{L}(\cdot)$ and the $k \times k$ matrix-valued process $Q(\cdot) := (Q_{ij}(\cdot))$ are called the pair-list process and the pair-type process, respectively. The terminal value $Q(T)$ of the latter is referred to as the *mating pattern* of the population. For the sake of convenience, we set $\mathcal{L}(t) := \mathcal{L}(T)$ and $Q(t) := Q(T)$ for every $t > T$.

Let ν be the probability measure that the firing times $\tau = (\tau_s(\zeta))_{s \geq 1, \zeta \in Z}$ induce on (Φ, \mathcal{F}) , where \mathcal{F} is the Borel σ -algebra corresponding to the product topology. Similarly, given any $\mathbf{t} \in \Phi$, let $\mathbb{P}^{\mathbf{t}}$ be the probability measure that (i) the sequence $((\Delta \mathcal{L})'(t_r^*))_{r \geq 1}$ of temporary pairs and (ii) the pair-list process $\mathcal{L}(\cdot)$ induce on some appropriate measurable space (Ω, \mathcal{B}) that is common for all $\mathbf{t} \in \Phi$. We prefer not to explicitly define (Ω, \mathcal{B}) since we will never directly refer to it. Finally, define \mathbb{P} as the semi-direct product measure of ν and $\mathbb{P}^{\mathbf{t}}$ on $(\Phi \times \Omega, \mathcal{F} \otimes \mathcal{B})$, i.e., $d\mathbb{P}(\mathbf{t}, \omega) := d\nu(\mathbf{t})d\mathbb{P}^{\mathbf{t}}(\omega)$. Note that $\mathbb{P}^{\mathbf{t}}(\cdot) = \mathbb{P}(\cdot | \tau = \mathbf{t})$ for $\mathbf{t} \in \text{supp}(\nu)$. We write $\mathbb{E}^{\mathbf{t}}$ (resp. \mathbb{E}) to denote expectation under $\mathbb{P}^{\mathbf{t}}$ (resp. \mathbb{P}).

The firing time distributions and the mating preference matrix are jointly referred to as the *encounter-mating (EM) law*. In this paper, we will consider the following EM laws:

- (Def) general firing times & definite mating upon encounter (Section 2);
- (Poi) Poisson firing times & general mating preferences (Section 3); and
- (Ber) Bernoulli firing times & general mating preferences (Section 4).

See Subsection 1.2 for the definitions of the quantities and concepts that are of central interest to us in this paper. As we outline in Subsection 1.3, cases (Poi) and (Ber) are generalizations of models that have been previously introduced in the literature, whereas our main purpose in analyzing case (Def) is to unify and clarify the other two. See Subsection 1.4 for a summary of our results.

1.2. Panmixia, homogamy and heterogamy. Given the EM law of the species and a population with the number of animals of each sex and type satisfying (1.1), we would like to analyze the distribution of the pair-type process $Q(\cdot)$. In our analysis, for any $i, j \in \{1, \dots, k\}$, we will frequently refer to the following quantities:

$$(1.3) \quad u_{ij}(t; x_1, \dots, x_k; y_1, \dots, y_k) := \mathbb{E}[Q_{ij}(t)] \quad \text{for any } t \geq 0;$$

$$(1.4) \quad u_{ij}^*(x_1, \dots, x_k; y_1, \dots, y_k) := \mathbb{E}[Q_{ij}(T)] \quad \text{for the terminal time } T; \text{ and}$$

$$(1.5) \quad u_{ij}^*(x_1, \dots, x_k; y_1, \dots, y_k | \mathbf{t}) := \mathbb{E}^{\mathbf{t}}[Q_{ij}(T)] \quad \text{for any } \mathbf{t} \in \Phi.$$

Among these, (1.4) will play a pivotal role.

We start with two observations. First, since all the animals mate by the terminal time T , the mating pattern $Q(T)$ is a random $k \times k$ contingency table whose i th row sum and j th column sum are equal to x_i and y_j , respectively. In particular, it has

$$(1.6) \quad k \cdot k - (k + k) + 1 = (k - 1)^2$$

degrees of freedom. Second, if there is definite mating upon encounter, i.e., if $p_{ij} = 1$ for every $i, j \in \{1, \dots, k\}$, then all of the $n!$ possible terminal pair-lists should be equally likely. (We will prove this later.) In particular, each of the x_i type- i females forms a permanent pair with a type- j male with probability y_j/n . This motivates the following definition.

DEFINITION 1.2. *A population is said to be panmictic if*

$$u_{ij}^*(x_1, \dots, x_k; y_1, \dots, y_k) = \frac{x_i y_j}{n}$$

for every $i, j \in \{1, \dots, k\}$, where $x_i \geq 0$ (resp. $y_j \geq 0$) is the number of type- i females (resp. type- j males), satisfying (1.1). The species is said to be panmictic if every population of animals from the species is panmictic.

In words, panmixia for a species refers to having (on average) zero correlation between female and male types in permanent pairs in any population.

To complement the concept of panmixia, homogamy (resp. heterogamy) is defined as females and males of similar types having positive (resp. negative) correlation in $\mathbb{E}[Q(T)]$. However,

in order to make these definitions precise, one has to choose an appropriate metric on the set $\{1, \dots, k\}$ of types. Also, note that the definitions of panmixia, homogamy and heterogamy are not a-priori collectively exhaustive even for a fixed population. For example, type-1 females and males can be positively correlated while type-2 females and males are negatively correlated. However, when $k = 2$, i.e., there are only two types of females and males, these potential issues are ruled out. Indeed, there is a unique choice of metric on the set $\{1, 2\}$, and $Q(T)$ has only one degree of freedom by (1.6). In particular, finding $u_{11}^*(x_1, x_2; y_1, y_2)$ is sufficient for determining the other three. Therefore, the following concise definitions make sense.

DEFINITION 1.3. *For $k = 2$, a population is said to be*

$$\begin{aligned} \text{homogamous if } u_{11}^*(x_1, x_2; y_1, y_2) &> \frac{x_1 y_1}{n}, \quad \text{and} \\ \text{heterogamous if } u_{11}^*(x_1, x_2; y_1, y_2) &< \frac{x_1 y_1}{n}, \end{aligned}$$

where $x_i \geq 1$ (resp. $y_j \geq 1$) is the number of type- i females (resp. type- j males), such that $n = x_1 + x_2 = y_1 + y_2$. The species is said to be homogamous (resp. heterogamous) if every population of animals from the species is homogamous (resp. heterogamous).

REMARK 1.4. In the definition above, we have not allowed $x_1 x_2 y_1 y_2$ to be 0 because, in that case, the mating pattern $Q(T)$ is deterministic and we trivially have

$$u_{11}^*(x_1, x_2; y_1, y_2) = \frac{x_1 y_1}{n}.$$

1.3. Previous results and related literature. In a 1988 paper, Gimelfarb [13] introduces two models of encounter-mating (EM) for permanent monogamous pair formation. Both of these models are in discrete time, and they differ from each other only in the firing times of the animals. In the first one, called individual EM, exactly one uniformly sampled single male (and no female) fires at each $t \in \mathbb{N}$, whereas in the second one, called mass EM, all single males (and no females) fire at each $t \in \mathbb{N}$. Given the firing times, the encounter and mating stages at each firing round of these models are as described in Subsection 1.1. Therefore, they are special cases of our SEM model. Indeed, for individual EM, the pair-type process is the discrete-time process embedded in $Q(\cdot)$ when $\{N(\zeta)\}_{\zeta \in Z}$ are Poisson processes with intensity 0 (resp. 1) for each $\zeta \in Z_f$ (resp. $\zeta \in Z_m$). Similarly, mass EM corresponds to the SEM model when $\{N(\zeta)\}_{\zeta \in Z}$ are Bernoulli processes with success probability 0 (resp. 1) for each $\zeta \in Z_f$ (resp. $\zeta \in Z_m$).

In order to simplify the analysis, Gimelfarb replaces all of the quantities such as $Q(\cdot)$ with their expectations. He says that, because of the law of large numbers (LLN), this is a good approximation when n is large, but he does not rigorously justify this claim. He defines the concepts of panmixia, homogamy and heterogamy, but only in an asymptotic sense as $n \rightarrow \infty$. He then asserts that the species is (asymptotically) panmictic whenever

$$(1.7) \quad p_{ij} = \bar{\alpha}_i + \bar{\beta}_j \quad \text{for some } \bar{\alpha}_1, \dots, \bar{\alpha}_k, \bar{\beta}_1, \dots, \bar{\beta}_k \text{ in the individual EM model, and}$$

$$(1.8) \quad p_{ij} = 1 - \bar{\gamma}_i \bar{\delta}_j \quad \text{for some } \bar{\gamma}_1, \dots, \bar{\gamma}_k, \bar{\delta}_1, \dots, \bar{\delta}_k \text{ in the mass EM model.}$$

He says that he was unable to prove that (1.7) is a sufficient condition for panmixia, and instead provides numerical evidence to back up this claim. In contrast, he does give an argument to show that (1.8) is a sufficient condition for panmixia. However, this argument is not rigorous because of his underlying LLN approximation (i.e., replacing quantities with their expectations).

Gimelfarb draws two main conclusions from the sufficient conditions (1.7) and (1.8) for panmixia. First, for either model, given a population, there is a many-to-one correspondence between mating preference matrices and expected mating patterns. This is a very important point; not only theoretically, but also practically. Indeed, the mating pattern of a population can be observed in its habitat whereas the mating preferences have to be determined by laboratory experiments which are relatively costly. It is therefore tempting to try to infer the latter from the former, but Gimelfarb concludes that this inverse problem cannot be solved. His second

conclusion is that knowing just the mating preference matrix is not enough for predicting the mating pattern. We need to also know whether we have, say, individual EM or mass EM. On a related note, Gimelfarb mentions that these two are of course not the only possible models. They are the extreme ones in some sense, and various intermediate cases can be considered, too. We will come back to this point later in Section 5.

Gimelfarb was not the first author to propose an earlier version of the SEM model. Indeed, what he calls individual EM was previously introduced by Romney [21] in the context of anthropology to model marriages in a community. Mosteller [20] analyzed Romney's model and gave recursive equations for the expected mating pattern, which can be iteratively solved when n is sufficiently small. What is particularly important about Mosteller's approach is that he did not make the LLN approximation of Gimelfarb, as the latter replaces the SEM model with a deterministic one. On this note, prior to Gimelfarb, such a deterministic EM model was proposed by Taylor [22] who gave a system of ODEs describing the evolution of $Q(t)$. However, he could provide only numerical solutions for this system. Also, neither Mosteller nor Taylor mentioned panmixia in their papers.

Panmixia is one of the fundamental concepts in population genetics. In particular, it is one of the main assumptions underlying the Hardy-Weinberg law which states that genotype frequencies remain constant in a population to which no evolutionary force acts on, see, e.g., [10] for details. In the literature, panmixia is also referred to as random mating. Gimelfarb [13] points out that the latter term is rather misleading especially for a bottom-up approach (i.e., from mating preferences to mating patterns) such as in the SEM model. Indeed, random mating suggests that the animals do not have any preferences, i.e., there exists a constant $p \in (0, 1]$ such that $p_{ij} = p$ for every $i, j \in \{1, \dots, k\}$. However, as we have already said, panmixia is about the expected mating pattern and there are many mating preference matrices that give rise to it.

In the cases of homogamy and heterogamy, the genotype frequencies might differ greatly from the ones predicted by the Hardy-Weinberg law, see Chapter 4 of [11] and the references therein. Also, for the effect of homogamy and heterogamy on the genetical evolution of a finite population, see e.g., [8], which is inspired by [16]. In these works, the terms positive and negative assortative mating are used to mean homogamy and heterogamy, respectively. Again, one should keep in mind that these terms refer to the expected mating pattern and not to the mating preferences. Indeed, (say) type-1 females in a homogamous population need not necessarily prefer mating with type-1 males. See Section 5 for a class of examples.

In the population dynamics literature, most models of pair formation assume that the females unilaterally accept or reject the males. This assumption is generally realistic (see [3] for the empirical aspects of sexual selection), and its various consequences have been studied in, e.g., [17], [18], and recently in [9]. Therefore, on one hand, the fact that the SEM model removes this assumption might seem unnecessary. On the other hand, not specifying which sex accepts or rejects the other one extends the scope of the model and makes it potentially suitable for two-sided matching problems that also have applications outside of biology. Such problems are typically studied using game theory, see the survey paper [4] and the series of papers [1, 2] regarding assortative mating. Moreover, in contrast to Gimelfarb's EM models, allowing both females and males to fire at prescribed rates not only makes the SEM model more versatile, but also introduces degrees of freedom in the EM law which then can be exploited to yield exact formulas under certain conditions. This is precisely what we will do in our analysis.

Having mentioned some of the advantages of the SEM model, we should point out that it clearly has limitations, too, as it is a model for permanent monogamous pair formation without births, deaths or offsprings. In his aforementioned paper [13], Gimelfarb also considers models where either one or both sexes are allowed to be polygamous. These models turn out to be much easier to analyze and less interesting in their behavior. Moreover, in another paper [12], he introduces a simple model of pair formation that allows separation of pairs and analyzes it using information theoretical concepts. Regarding models incorporating births, deaths and

offsprings, we refer the reader to the recent paper by Haderler [15] and the references therein. From a biological point of view, the fact that our model does not include these factors can be partially justified by assuming that everything takes place in one breeding season.

Finally, as the aforementioned work of Romney [21] attests, the scope of the SEM model is not restricted to non-human animals. Indeed, the demographic, cultural, and technological changes of the last 10,000 years did not preclude the potential for natural and sexual selection in our species [5]. Determining the mating patterns of human populations has crucial applications, e.g., modelling how sexually transmitted diseases spread, see [7]. However, we can attempt to answer such questions only after incorporating separations and/or polygamy into our model.

1.4. Summary of our results. Before giving the precise statements of our results on the SEM model in Sections 2, 3 and 4, which require technical assumptions and further notation, we will summarize them below. While doing so, we will provide motivation for them, and explain how they are related to each other as well as to the previous results in the literature.

1.4.1. General firing times. Recall that, while describing the SEM model in Subsection 1.1, we have made certain assumptions regarding the joint distribution of the point processes $\{N(\zeta)\}_{\zeta \in \mathbb{Z}}$ which give the firing times $\tau = (\tau_s(\zeta))_{s \geq 1}$ of the animals. There, we have also said that we will make further technical assumptions to ensure that $\tau \in \Phi$ (see Definition 1.1) almost surely. We start Section 2 by listing all of these assumptions which hold throughout the paper.

The sampling procedure at the encounter stage (see Subsection 1.1) is the most complicated aspect of the SEM model. In Subsection 2.2, we give an alternative representation of the model which amounts to decomposing the encounter stage at any firing round into two steps: at the first one, we pair up all of the single animals uniformly and without replacement; and at the second one, we discard the pairs that do not have any animals that have fired in that round. We make repeated use of this representation in our arguments.

In Subsection 2.3, we consider the SEM model under the assumption of definite mating upon encounter, i.e., we take $p_{ij} = 1$ for every $i, j \in \{1, \dots, k\}$. In this special case, it is possible to give a detailed analysis and provide exact formulas for (1.3), (1.4) and (1.5). This is due to the fact that the mating stages at the firing rounds are trivial, i.e., the two-stage structure of the model is reduced to one. Our first result is Theorem 2.3 which says that the terminal pair-list $\mathcal{L}(T)$ is uniformly distributed under $\mathbb{P}^{\mathbf{t}}$ for every $\mathbf{t} \in \Phi$. In particular, $\mathcal{L}(T)$ and $\tau = (\tau_s(\zeta))_{s \geq 1}$ are independent under \mathbb{P} . This is very intuitive. Indeed, no animal rejects the member of the opposite sex that it is randomly paired with, and therefore the order with which the pairs are formed should not matter. This elementary result has many important consequences. First of all, since the mating pattern $Q(T)$ is a function of the terminal pair-list $\mathcal{L}(T)$, the distribution of $Q(T)$ under $\mathbb{P}^{\mathbf{t}}$ for every $\mathbf{t} \in \Phi$ can be easily computed and turns out to be multiple hypergeometric. Therefore, the expected mating pattern under $\mathbb{P}^{\mathbf{t}}$ is a contingency table in product form for every $\mathbf{t} \in \Phi$. In particular, the species is panmictic as one would predict.

Our main result in the case of definite mating upon encounter is Theorem 2.6 which gives the distribution of $Q(t)$ under \mathbb{P} for any $t \geq 0$ (rather than just the terminal time T). This result is yet another consequence of Theorem 2.3. Indeed, the proof of Theorem 2.6 uses the alternative representation of the model, and relies on the observation that the pair-list process $\mathcal{L}(\cdot)$ is measurable w.r.t. $\mathcal{L}(T)$ and $\tau = (\tau_s(\zeta))_{s \geq 1}$ which are independent under \mathbb{P} (as shown in Theorem 2.3). Finally, for any $t \geq 0$, once the distribution of $Q(t)$ under \mathbb{P} is known, the expected pair-type matrix $\mathbb{E}[Q(t)]$ is easily computed.

We are able to obtain all of these exact formulas in the special case of definite mating upon encounter, but this is unfortunately not representative of the generic case. In Subsection 2.4, as a first step in dealing with nontrivial mating preferences, we consider the SEM model under the assumption that, for every $i, j \in \{1, \dots, k\}$, $p_{ij} = 1$ holds if and only if $i \neq j$. In words, there is definite mating upon encounter for mixed-type temporary pairs only. Here and throughout, pure-type (resp. mixed-type) refers to animals in a pair having the same (resp. different) types.

In this case, the expectation of $Q_{ii}(T)$ under \mathbb{P} for pure-type pairs should be strictly less than what it was in the previous case. This prediction turns out to be true even under $\mathbb{P}^{\mathbf{t}}$ for every $\mathbf{t} \in \Phi$, and its precise statement constitutes Theorem 2.10.

At first sight, the special cases of mating preferences considered in Subsections 2.3 and 2.4 seem to be limited in their scope. However, as we will outline below, when the point processes $\{N(\zeta)\}_{\zeta \in Z}$ are Poisson or Bernoulli, it is possible to generalize our results for definite mating upon encounter to a wide class of mating preferences.

1.4.2. Poisson firing times. In Section 3, we consider the SEM model under the assumption that the point processes $\{N(\zeta)\}_{\zeta \in Z}$ are Poisson. For any $\zeta \in Z$, we denote the intensity of $N(\zeta)$ by α_i (resp. β_j) if ζ is a type- i female (resp. type- j male). In this case, it is clear that $Q(\cdot)$ is a continuous-time pure jump Markov chain under \mathbb{P} . Moreover, its entries have jumps of size exactly 1 since there are no multiple firings at any time. In Subsection 3.1, we give the infinitesimal generator of $Q(\cdot)$ which turns out to depend on p_{ij} , α_i and β_j only through

$$(1.9) \quad \pi_{ij} := p_{ij}(\alpha_i + \beta_j).$$

Therefore, we have the freedom to change these parameters as long as the $k \times k$ matrix

$$\Pi := (\pi_{ij})$$

stays the same. This observation plays a key role in our analysis. We finish Subsection 3.1 by giving recursive equations for $u_{ij}(t; x_1, \dots, x_k; y_1, \dots, y_k)$ and $u_{ij}^*(x_1, \dots, x_k; y_1, \dots, y_k)$.

Next, in Subsection 3.2, we introduce the so-called Poisson fine balance condition on Π which requires that $\pi_{ij} + \pi_{i'j'} = \pi_{ij'} + \pi_{i'j}$ for every $i, j, i', j' \in \{1, \dots, k\}$. We motivate this condition by showing that it is necessary for panmixia. Then, we prove that the Poisson fine balance condition holds if and only if there exist $\bar{\alpha}_1, \dots, \bar{\alpha}_k, \bar{\beta}_1, \dots, \bar{\beta}_k \geq 0$ such that

$$(1.10) \quad \pi_{ij} = \bar{\alpha}_i + \bar{\beta}_j$$

for every $i, j \in \{1, \dots, k\}$.

Our main result in Section 3 is Theorem 3.6 which assumes that the Poisson fine balance condition is satisfied, and gives the distribution of $Q(t)$ under \mathbb{P} for every $t \geq 0$. The proof of Theorem 3.6 uses the aforementioned change-of-parameters technique. Indeed, by comparing (1.9) and (1.10), we can WLOG assume that

$$p_{ij} = 1, \quad \alpha_i = \bar{\alpha}_i \quad \text{and} \quad \beta_j = \bar{\beta}_j$$

for every $i, j \in \{1, \dots, k\}$. Then, we have definite mating upon encounter and Theorem 2.6 gives the desired result. The corollaries of Theorem 2.6 carry over, too. In particular, the Poisson fine balance condition characterizes panmixia in the Poisson case.

As we have mentioned in Subsection 1.3, Gimelfarb's individual EM model corresponds to having Poisson firing times with intensities $\alpha_i = 0$ and $\beta_j = 1$ for every $i, j \in \{1, \dots, k\}$. Recall condition (1.7) which Gimelfarb conjectures to be sufficient for (asymptotic) panmixia. Note that, if (1.7) holds for some $\bar{\alpha}_1, \dots, \bar{\alpha}_k, \bar{\beta}_1, \dots, \bar{\beta}_k$, then

$$\pi_{ij} = p_{ij}(\alpha_i + \beta_j) = (\bar{\alpha}_i + \bar{\beta}_j)(0 + 1) = \bar{\alpha}_i + \bar{\beta}_j,$$

i.e., the Poisson fine balance condition is satisfied by (1.10). Therefore, the species is indeed panmictic. At this point, we would like to emphasize that we thereby not only settle Gimelfarb's panmixia conjecture, but also strengthen and generalize it in the following ways:

- We prove that the Poisson fine balance condition is sufficient for panmixia (and not only for asymptotic panmixia).
- We show that, in fact, the Poisson fine balance condition characterizes panmixia, i.e., it is also necessary.
- We allow the intensities $\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_k$ of the Poisson firing times to be arbitrary.
- When the Poisson fine balance condition is satisfied, we give an explicit formula for the distribution of $Q(t)$ under \mathbb{P} for any $t \geq 0$.

Finally, in Subsection 3.3, we assume that $k = 2$, i.e., there are only two types of females and males. Now that panmixia is characterized by the Poisson fine balance condition, it is natural to ask if homogamy and heterogamy can be similarly characterized. We accomplish this in Theorem 3.9 which says that the species is

$$(1.11) \quad \begin{aligned} &\text{heterogamous if } \pi_{11} + \pi_{22} < \pi_{12} + \pi_{21}, \\ &\text{panmictic if } \pi_{11} + \pi_{22} = \pi_{12} + \pi_{21}, \quad \text{and} \\ &\text{homogamous if } \pi_{11} + \pi_{22} > \pi_{12} + \pi_{21}. \end{aligned}$$

The proof of Theorem 3.9 uses the change-of-parameters technique, too. Indeed, e.g., when $\pi_{11} + \pi_{22} < \pi_{12} + \pi_{21}$, we can WLOG assume that we have definite mating upon encounter for mixed-type temporary pairs only, and the desired result follows from Theorem 2.10. The significance of Theorem 3.9 lies in the fact that it provides a full characterization of panmixia, homogamy and heterogamy. Every Π satisfies exactly one of the three conditions in (1.11), i.e., we have a trichotomy.

1.4.3. Bernoulli firing times. In Section 4, we consider the SEM model under the assumption that $\{N(\zeta)\}_{\zeta \in Z}$ are Bernoulli point processes. This means that, for any $\zeta \in Z$, $N(\zeta)$ is a random subset of \mathbb{N} formed by independent Bernoulli trials assigned to the natural numbers. We denote the success probability of these Bernoulli trials by α_i (resp. β_j) if ζ is a type- i female (resp. type- j male).

Our results in the case with Bernoulli firing times are very similar to those in the previous case with Poisson firing times. In order to emphasize this similarity, we have chosen to use the same wording (wherever possible) and parallel numbering for the theorems and displays in Sections 3 and 4. This way, these two sections can be read independently of each other, and their contents can be easily compared and contrasted.

Having said this, here are some of the important points where the Bernoulli case differs from the Poisson case.

- The pair-type process $Q(\cdot)$ is a discrete-time (as opposed to continuous-time) Markov chain under \mathbb{P} . The entries of $Q(\cdot)$ can have integer jumps of size greater than 1 since multiple firings are possible at any time $t \in \mathbb{N}$.
- The transition kernel of $Q(\cdot)$ depends on p_{ij} , α_i and β_j only through

$$\pi_{ij} := p_{ij}(\alpha_i + \beta_j - \alpha_i\beta_j),$$

and therefore, we have the freedom to change these parameters as long as $\Pi := (\pi_{ij})$ is the same.

- The Poisson fine balance condition is replaced by the so-called Bernoulli fine balance condition which requires that

$$(1 - \pi_{ij})(1 - \pi_{i'j'}) = (1 - \pi_{ij'})(1 - \pi_{i'j})$$

for every $i, j, i', j' \in \{1, \dots, k\}$.

- The Bernoulli fine balance condition holds iff there exist $\bar{\alpha}_1, \dots, \bar{\alpha}_k, \bar{\beta}_1, \dots, \bar{\beta}_k \in [0, 1]$ such that

$$1 - \pi_{ij} = (1 - \bar{\alpha}_i)(1 - \bar{\beta}_j)$$

for every $i, j \in \{1, \dots, k\}$.

Our main result in Section 4 is Theorem 4.6 which is completely parallel to Theorem 3.6. Its proof uses the change-of-parameters idea and thus reduces the model to definite mating upon encounter. The distributions that Theorems 3.6 and 4.6 provide for $Q(t)$ under \mathbb{P} are almost identical, except for the natural difference due to continuous-time vs. discrete-time. In particular, the Bernoulli fine balance condition characterizes panmixia in the Bernoulli case.

Recall from Subsection 1.3 that Gimelfarb's mass EM model corresponds to having Bernoulli firing times with success probabilities $\alpha_i = 0$ and $\beta_j = 1$ for every $i, j \in \{1, \dots, k\}$. His sufficient

condition (1.8) for panmixia implies the Bernoulli fine balance condition. Indeed, if (1.8) holds for some $\bar{\gamma}_1, \dots, \bar{\gamma}_k, \bar{\delta}_1, \dots, \bar{\delta}_k$, then

$$1 - \pi_{ij} = 1 - p_{ij}(\alpha_i + \beta_j - \alpha_i \beta_j) = 1 - (1 - \bar{\gamma}_i \bar{\delta}_j)(0 + 1 - 0 \cdot 1) = \bar{\gamma}_i \bar{\delta}_j = (1 - \bar{\alpha}_i)(1 - \bar{\beta}_j)$$

with $\bar{\alpha}_i = 1 - \bar{\gamma}_i$ and $\bar{\beta}_j = 1 - \bar{\delta}_j$. Therefore, we not only provide a rigorous proof of Gimelfarb's panmixia result, but also strengthen and generalize it as in the previous case.

Finally, we assume that $k = 2$, and prove in Theorem 4.9 that the species is

$$\begin{aligned} &\text{heterogamous if } (1 - \pi_{11})(1 - \pi_{22}) > (1 - \pi_{12})(1 - \pi_{21}), \\ &\text{panmictic if } (1 - \pi_{11})(1 - \pi_{22}) = (1 - \pi_{12})(1 - \pi_{21}), \text{ and} \\ &\text{homogamous if } (1 - \pi_{11})(1 - \pi_{22}) < (1 - \pi_{12})(1 - \pi_{21}). \end{aligned}$$

Note that this trichotomy is completely parallel to (1.11) from the previous case.

2. GENERAL FIRING TIMES

2.1. Assumptions and further notation. In this section, we will make the following rather general assumptions regarding the firing times of the animals.

- (Gen1) $\{N(\zeta)\}_{\zeta \in Z}$ are mutually independent simple point processes on $[0, \infty)$.
- (Gen2) $N(\zeta)$ are identically distributed for all type- i females (resp. type- j males).
- (Gen3) $N(\zeta)\{0\}$ is almost surely zero for every $\zeta \in Z$.
- (Gen4) $N(\zeta)[0, \infty)$ is almost surely either zero or infinite for each $\zeta \in Z$. Moreover, if it is zero for a female (resp. male), then it is infinite for all males (resp. females).

Explanation: (Gen1) and (Gen2) were already mentioned in Subsection 1.1. (Gen3) is equivalent to saying that the firing times are almost surely positive, which implies that $\mathcal{L}(0) = \emptyset$. Finally, on one hand, (Gen4) allows certain animals to not fire at all; on the other hand, in combination with the $p_{ij} > 0$ assumption, it ensures that such animals eventually mate (upon an encounter initiated by a member of the opposite sex). Recall Definition 1.1 and note that (Gen1)–(Gen4) imply $\tau \in \Phi$ almost surely. Here, as a convention, if $N(\zeta)$ has no time points, i.e., $N(\zeta)[0, \infty) = \emptyset$, we set $\tau_s(\zeta) = \infty$ for every $s \geq 1$.

Let $\mathcal{M}^{k \times k}(\mathbb{N} \cup \{0\})$ be the set of $k \times k$ matrices with nonnegative integer entries, equipped with the following partial order: $M \leq M'$ iff $m_{ij} \leq m'_{ij}$ for every $i, j \in \{1, \dots, k\}$. Denote the i th row sum, the j th column sum and the grand total of any $M = (m_{ij}) \in \mathcal{M}^{k \times k}(\mathbb{N} \cup \{0\})$ by

$$m_{i\cdot} = \sum_{j'=1}^k m_{ij'}, \quad m_{\cdot,j} = \sum_{i'=1}^k m_{i'j} \quad \text{and} \quad m_{tot} = \sum_{i'=1}^k \sum_{j'=1}^k m_{i'j'},$$

respectively. With this notation, the pair-type process $Q(\cdot)$ takes values in

$$(2.1) \quad \mathcal{E} = \mathcal{E}(x_1, \dots, x_k; y_1, \dots, y_k) = \{M \in \mathcal{M}^{k \times k}(\mathbb{N} \cup \{0\}) : m_{i\cdot} \leq x_i \text{ and } m_{\cdot,j} \leq y_j\}$$

and its initial value is $Q(0) = 0$, the zero matrix.

At any time $t \in [0, T]$, the number of single type- i females, the number of single type- j males and the total number of single females (or males) are equal to

$$x_i - Q_{i\cdot}(t), \quad y_j - Q_{\cdot,j}(t) \quad \text{and} \quad n - Q_{tot}(t),$$

respectively. In particular, the mating pattern $Q(T)$ takes values in the set

$$\mathcal{E}' = \mathcal{E}'(x_1, \dots, x_k; y_1, \dots, y_k) = \{M \in \mathcal{M}^{k \times k}(\mathbb{N} \cup \{0\}) : m_{i\cdot} = x_i \text{ and } m_{\cdot,j} = y_j\}$$

of all $k \times k$ contingency tables with i th row sum (resp. j th column sum) equal to x_i (resp. y_j). Note that $|\mathcal{E}'| > 1$ if and only if

$$|\{i = 1, \dots, k : x_i > 0\}| > 1 \quad \text{and} \quad |\{j = 1, \dots, k : y_j > 0\}| > 1.$$

It is clear that the pair-type matrix $Q(t)$ at any $t \geq 0$ is measurable w.r.t. the pair-list $\mathcal{L}(t)$. Indeed, define a function $\gamma_1 : \{\varphi_1, \dots, \varphi_n\} \rightarrow \{1, \dots, k\}$ by setting $\gamma_1(\varphi_a) = i$ if φ_a is of type- i .

Similarly, define a function $\gamma_2 : \{\mu_1, \dots, \mu_n\} \rightarrow \{1, \dots, k\}$ which encodes the type of each male. Finally, let $\Gamma(\varphi_a, \mu_b) := (\gamma_1(\varphi_a), \gamma_2(\mu_b))$. With this notation,

$$(2.2) \quad \begin{aligned} x_i &= \sum_{a=1}^n \mathbb{I}_{\{\gamma_1(\varphi_a)=i\}}, & y_j &= \sum_{b=1}^n \mathbb{I}_{\{\gamma_2(\mu_b)=j\}} \quad \text{and} \\ Q_{ij}(t) &= \sum_{a=1}^n \sum_{b=1}^n \mathbb{I}_{\{(\varphi_a, \mu_b) \in \mathcal{L}(t), \Gamma(\varphi_a, \mu_b)=(i,j)\}}. \end{aligned}$$

Observe that the terminal pair-list $\mathcal{L}(T)$ consists of exactly n pairs, i.e., $|\mathcal{L}(T)| = n$, and that no two pairs in it have a common animal. We will refer to the latter property as being admissible. Each realization of $\mathcal{L}(T)$ can be identified with an element σ of the symmetric group Σ_n (i.e., a permutation of $\{1, \dots, n\}$) in the following way:

$$(2.3) \quad \sigma(a) = b \quad \Longleftrightarrow \quad (\varphi_a, \mu_b) \in \mathcal{L}(T).$$

We will abbreviate this identification as $\mathcal{L}(T) = \mathcal{C}_\sigma$. With this notation, $\mathcal{L}(T)$ is sampled from

$$\Lambda_n := \{\mathcal{C}_\sigma : \sigma \in \Sigma_n\}.$$

2.2. An alternative representation of the SEM model. Recall (1.2) and the sampling procedure at the encounter stage of the first firing round. The following lemma gives the probability measure this procedure induces on the set of admissible collections of pairs.

LEMMA 2.1. *Given any $\mathbf{t} \in \Phi$, fix an admissible collection \mathcal{C} of pairs such that*

- (i) *each $\zeta \in Z$ with $t_1(\zeta) = t_1^*$ is in a pair in \mathcal{C} ; and*
- (ii) *each pair $(\varphi_a, \mu_b) \in \mathcal{C}$ satisfies $t_1(\varphi_a) \wedge t_1(\mu_b) = t_1^*$.*

Then,

$$(2.4) \quad \mathbb{P}^{\mathbf{t}}((\Delta\mathcal{L})'(t_1^*) = \mathcal{C}) = \frac{1}{n(n-1) \cdots (n-|\mathcal{C}|+1)} = \frac{(n-|\mathcal{C}|)!}{n!}.$$

Proof. We will show this by induction on $n \geq 1$. If $n = 1$, then there is only one possible collection $\mathcal{C} := \{(\varphi_1, \mu_1)\}$, which is consistent with

$$\frac{(n-|\mathcal{C}|)!}{n!} = \frac{(1-1)!}{1!} = 1.$$

For any $n \geq 2$, assume that the desired result holds for $n-1$. The probability that the first animal (w.r.t. the order on $\{\zeta \in Z : t_1(\zeta) = t_1^*\}$), say φ_a , indeed samples whoever it is paired up with in \mathcal{C} , say μ_b , is equal to $\frac{1}{n}$. If $|\mathcal{C}| = 1$, then we are done; if not, then we have the following left: (i) $n-1$ females and males for the subsequent samplings at t_1^* ; and (ii) $|\mathcal{C}| - 1$ pairs in $\mathcal{C} \setminus \{(\varphi_a, \mu_b)\}$. Therefore, the sought probability is equal to the RHS of (2.4) by the induction hypothesis. \square

REMARK 2.2. It is evident from this proof that the distribution of $(\Delta\mathcal{L})'(t_1^*)$ under $\mathbb{P}^{\mathbf{t}}$ does not depend on the order imposed on $\{\zeta \in Z : t_1(\zeta) = t_1^*\}$.

Observe that if we sample a $\sigma \in \Sigma_n$ uniformly at random and identify it with an admissible collection \mathcal{C}_σ of n pairs via (2.3), then the probability that it contains (as a subset) a given admissible collection \mathcal{C} is equal to the RHS of (2.4). This simple observation provides us with the following alternative representation of the SEM model: Right before the r th firing round starts, the pair-list is equal to $\mathcal{L}(t_{r-1}^*)$, and there are $n - Q_{tot}(t_{r-1}^*) \geq 1$ single females and males. At t_r^* , relabel these animals $1, \dots, n - Q_{tot}(t_{r-1}^*)$ and pair them up according to \mathcal{C}_σ where $\sigma \in \Sigma_{n-Q_{tot}(t_{r-1}^*)}$ is sampled uniformly at random. Next, form $(\Delta\mathcal{L})'(t_r^*)$ by temporarily keeping only those pairs that have at least one animal that fires at the r th round. Finally, for each such type- ij temporary pair, sample a Bernoulli random variable with success probability p_{ij} and discard the pair in the event of failure. This gives us the set $\Delta\mathcal{L}(t_r^*)$ of new permanent pairs which are then added to $\mathcal{L}(t_{r-1}^*)$ to form the pair-list $\mathcal{L}(t_r^*) = \mathcal{L}(t_{r-1}^*) \cup \Delta\mathcal{L}(t_r^*)$.

Although the rather top-down encounter mechanism in this representation of the SEM model is perhaps not natural from a biological point of view, it will turn out to be very convenient for mathematical analysis.

2.3. Definite mating upon encounter. Assume that $p_{ij} = 1$ for every $i, j \in \{1, \dots, k\}$. Then, the two stages of the SEM model are reduced to one since every encounter definitely results in mating. In particular, each animal mates by its first firing time at the latest, and therefore its subsequent firing times are irrelevant. In this case, it is intuitively clear that the terminal pair-list $\mathcal{L}(T)$ should be uniformly distributed on $\Lambda_n = \{\mathcal{C}_\sigma : \sigma \in \Sigma_n\}$ under \mathbb{P} . In fact, we have a much stronger result.

THEOREM 2.3. *If $p_{ij} = 1$ for every $i, j \in \{1, \dots, k\}$, then the terminal pair-list $\mathcal{L}(T)$ is uniformly distributed on Λ_n under $\mathbb{P}^{\mathbf{t}}$ for every $\mathbf{t} \in \Phi$. In particular, $\mathcal{L}(T)$ and $\tau = (\tau_s(\zeta))_{s \geq 1, \zeta \in Z}$ are independent under \mathbb{P} .*

Proof. We will prove this by strong induction on $n \geq 1$. The case $n = 1$ is trivial since there is only one possible collection $\{(\varphi_1, \mu_1)\}$ of pairs. For any $n \geq 2$, assume that the desired result holds for all $1 \leq n' \leq n - 1$. Fix a $\sigma \in \Sigma_n$ and let

$$\mathcal{C}_\sigma(t_1^*) = \{(\varphi_a, \mu_b) \in \mathcal{C}_\sigma : \tau_1(\varphi_a) \wedge \tau_1(\mu_b) = t_1^*\}.$$

Since $\mathcal{L}(0) = \emptyset$ and every encounter results in mating, we have $\mathcal{L}(t_1^*) = \Delta \mathcal{L}(t_1^*) = (\Delta \mathcal{L})'(t_1^*)$. Therefore,

$$\mathbb{P}^{\mathbf{t}}(\mathcal{L}(t_1^*) = \mathcal{C}_\sigma(t_1^*)) = \frac{(n - |\mathcal{C}_\sigma(t_1^*)|)!}{n!}$$

by Lemma 2.1 (or, equivalently, the alternative representation of the model). If $|\mathcal{C}_\sigma(t_1^*)| = n$, then $\mathcal{C}_\sigma = \mathcal{C}_\sigma(t_1^*)$, $T = t_1^*$, and we are done; if not, we have the following left: (i) $n - |\mathcal{C}_\sigma(t_1^*)|$ single females and males; and (ii) $n - |\mathcal{C}_\sigma(t_1^*)|$ pairs in $\mathcal{C}_\sigma \setminus \mathcal{C}_\sigma(t_1^*)$. Therefore,

$$\begin{aligned} \mathbb{P}^{\mathbf{t}}(\mathcal{L}(T) = \mathcal{C}_\sigma) &= \mathbb{P}^{\mathbf{t}}(\mathcal{L}(t_1^*) = \mathcal{C}_\sigma(t_1^*)) \mathbb{P}^{\mathbf{t}}(\mathcal{L}(T) = \mathcal{C}_\sigma \mid \mathcal{L}(t_1^*) = \mathcal{C}_\sigma(t_1^*)) \\ &= \frac{(n - |\mathcal{C}_\sigma(t_1^*)|)!}{n!} \frac{1}{(n - |\mathcal{C}_\sigma(t_1^*)|)!} = \frac{1}{n!} \end{aligned}$$

by the induction hypothesis. \square

Since the mating pattern $Q(T)$ is a function of the terminal pair-list $\mathcal{L}(T)$, we immediately obtain the following corollaries.

COROLLARY 2.4. *Take a population with x_1, \dots, x_k females and y_1, \dots, y_k males of types $1, \dots, k$, respectively, such that (1.1) holds. If $p_{ij} = 1$ for every $i, j \in \{1, \dots, k\}$, then*

$$\mathbb{P}^{\mathbf{t}}(Q(T) = M) = \frac{\left(\prod_i x_i!\right) \left(\prod_j y_j!\right)}{n! \left(\prod_{i,j} m_{ij}!\right)}$$

for every $\mathbf{t} \in \Phi$ and $M \in \mathcal{E}'$, i.e., the distribution of $Q(T)$ under $\mathbb{P}^{\mathbf{t}}$ (and, therefore, under \mathbb{P}) is multiple hypergeometric.

Proof. By Theorem 2.3, this is just a counting exercise. See [6, p. 247] for the multiple hypergeometric distribution which appears in the tests of randomness for contingency tables. Our desired result is stated on that page, too. Its proof is easy and we leave it to the reader. \square

COROLLARY 2.5. *If $p_{ij} = 1$ for every $i, j \in \{1, \dots, k\}$, then*

$$u_{ij}^*(x_1, \dots, x_k; y_1, \dots, y_k \mid \mathbf{t}) = \mathbb{E}^{\mathbf{t}}[Q_{ij}(T)] = \frac{x_i y_j}{n}$$

for every $\mathbf{t} \in \Phi$ and every $x_1, \dots, x_k, y_1, \dots, y_k \geq 0$ satisfying (1.1). Consequently, the species is panmictic.

Proof. By Corollary 2.4, this follows from the formula for the mean of a multiple hypergeometric distribution which is stated in [6, p. 247] and derived exactly in the same way as that of a hypergeometric distribution. \square

Under our current assumption of definite mating upon encounter, the pair-list process $\mathcal{L}(\cdot)$ is measurable w.r.t. the terminal pair-list $\mathcal{L}(T)$ and the firing times $\tau = (\tau_s(\zeta))_{s \geq 1, \zeta \in Z}$. Indeed, $\mathcal{L}(T)$ is the list of “who is destined to be with whom”, and

$$(2.5) \quad \mathcal{L}(t) = \{(\varphi_a, \mu_b) \in \mathcal{L}(T) : \tau_1(\varphi_a) \wedge \tau_1(\mu_b) \leq t\}$$

for $t \geq 0$. Due to the independence of $\mathcal{L}(T)$ and $(\tau_s(\zeta))_{s \geq 1, \zeta \in Z}$ (established in Theorem 2.3), this approach is not only conceptually elegant, but also computationally practical, as we will see in the theorem below.

Note that, naturally, only the first firing time of each animal appears in (2.5). Denote the cumulative distribution function (CDF) of these times by

$$F_i(t) := \nu(\tau_1(\varphi_a) \leq t) \quad \text{if } \gamma_1(\varphi_a) = i \quad \text{and} \quad G_j(t) := \nu(\tau_1(\mu_b) \leq t) \quad \text{if } \gamma_2(\mu_b) = j.$$

Since the pair-type process $Q(\cdot)$ is measurable w.r.t. $\mathcal{L}(\cdot)$ via (2.2), we have the following result.

THEOREM 2.6. *Take a population with x_1, \dots, x_k females and y_1, \dots, y_k males of types $1, \dots, k$, respectively, such that (1.1) holds. If $p_{ij} = 1$ for every $i, j \in \{1, \dots, k\}$, then*

$$\mathbb{P}(Q(t) = M) = \frac{\left(\prod_i x_i!\right) \left(\prod_j y_j!\right)}{n! \left(\prod_{i,j} m_{ij}!\right)} \sum_{\substack{M' \in \mathcal{E}' : \\ M' \geq M}} \prod_{i,j} \frac{(\lambda_{ij}(t))^{m_{ij}} (1 - \lambda_{ij}(t))^{m'_{ij} - m_{ij}}}{(m'_{ij} - m_{ij})!}$$

for every $t \geq 0$ and $M \in \mathcal{E}$, where

$$(2.6) \quad \lambda_{ij}(t) := 1 - (1 - F_i(t))(1 - G_j(t)) = F_i(t) + G_j(t) - F_i(t)G_j(t).$$

Proof. For every $t \geq 0$ and $M \in \mathcal{E}$, we have

$$(2.7) \quad \begin{aligned} \mathbb{P}(Q(t) = M) &= \sum_{\substack{M' \in \mathcal{E}' : \\ M' \geq M}} \mathbb{P}(Q(T) = M') \mathbb{P}(Q(t) = M \mid Q(T) = M') \\ &= \sum_{\substack{M' \in \mathcal{E}' : \\ M' \geq M}} \mathbb{P}(Q(T) = M') \prod_{i,j} \binom{m'_{ij}}{m_{ij}} (\lambda_{ij}(t))^{m_{ij}} (1 - \lambda_{ij}(t))^{m'_{ij} - m_{ij}} \end{aligned}$$

$$(2.8) \quad \begin{aligned} &= \sum_{\substack{M' \in \mathcal{E}' : \\ M' \geq M}} \frac{\left(\prod_i x_i!\right) \left(\prod_j y_j!\right)}{n! \left(\prod_{i,j} m'_{ij}!\right)} \prod_{i,j} \binom{m'_{ij}}{m_{ij}} (\lambda_{ij}(t))^{m_{ij}} (1 - \lambda_{ij}(t))^{m'_{ij} - m_{ij}} \\ &= \frac{\left(\prod_i x_i!\right) \left(\prod_j y_j!\right)}{n! \left(\prod_{i,j} m_{ij}!\right)} \sum_{\substack{M' \in \mathcal{E}' : \\ M' \geq M}} \prod_{i,j} \frac{(\lambda_{ij}(t))^{m_{ij}} (1 - \lambda_{ij}(t))^{m'_{ij} - m_{ij}}}{(m'_{ij} - m_{ij})!}. \end{aligned}$$

Explanation: $Q(T)$ and $(\tau_s(\zeta))_{s \geq 1, \zeta \in Z}$ are independent by Theorem 2.3. Recall (2.5) and note that each type- ij pair in $\mathcal{L}(T)$ is contained in $\mathcal{L}(t)$ with probability $\lambda_{ij}(t)$ and independently of all other pairs. Therefore, the conditional distribution of $Q_{ij}(t)$ given $Q_{ij}(T)$ is binomial with parameters $Q_{ij}(T)$ and $\lambda_{ij}(t)$. This gives (2.7). Finally, (2.8) follows from Corollary 2.4. \square

COROLLARY 2.7. *If $p_{ij} = 1$ for every $i, j \in \{1, \dots, k\}$, then*

$$u_{ij}(t; x_1, \dots, x_k; y_1, \dots, y_k) = \mathbb{E}[Q_{ij}(t)] = \frac{x_i y_j \lambda_{ij}(t)}{n}$$

for every $t \geq 0$ and $x_1, \dots, x_k, y_1, \dots, y_k \geq 0$ satisfying (1.1), where $\lambda_{ij}(t)$ is defined in (2.6).

Proof. We have seen in the proof of Theorem 2.6 that

$$Q_{ij}(t) \mid Q_{ij}(T) \sim B(Q_{ij}(T), \lambda_{ij}(t))$$

where B denotes the binomial distribution. Therefore,

$$\mathbb{E}[Q_{ij}(t)] = \mathbb{E}[\mathbb{E}[Q_{ij}(t) \mid Q_{ij}(T)]] = \mathbb{E}[Q_{ij}(T)]\lambda_{ij}(t) = \frac{x_i y_j \lambda_{ij}(t)}{n}$$

by the law of total expectation and Corollary 2.5. \square

2.4. Definite mating upon encounter for mixed-type temporary pairs only. Assume that, for every $i, j \in \{1, \dots, k\}$, $p_{ij} = 1$ holds if and only if $i \neq j$. In words, there is definite mating upon encounter for mixed-type temporary pairs only. In this case, it is intuitively clear that the expected number of pure-type pairs should be strictly less than what it was when we had definite mating upon encounter for all temporary pairs. This assertion turns out to be true not only under \mathbb{P} , but also under $\mathbb{P}^{\mathbf{t}}$ for every $\mathbf{t} \in \Phi$. We start with two simple observations.

LEMMA 2.8. For $x'_1, \dots, x'_k, y'_1, \dots, y'_k \geq 0$ with $n' := x'_1 + \dots + x'_k = y'_1 + \dots + y'_k$, define

$$(2.9) \quad H(x'_1, \dots, x'_k; y'_1, \dots, y'_k) := \begin{cases} \frac{x'_1 y'_1}{n'} & \text{if } n' \neq 0; \text{ and} \\ 0 & \text{if } n' = 0. \end{cases}$$

Then,

$$(2.10) \quad H(x'_1 - c_1, \dots, x'_k - c_k; y'_1 - c_1, \dots, y'_k - c_k) + c_1$$

is increasing in c_i for every $i \in \{1, \dots, k\}$. Moreover, if $x'_1 x'_2 y'_1 y'_2 \neq 0$, then (2.10) is strictly increasing in c_1 (resp. c_2) when $c_2 < x'_2 \wedge y'_2$ (resp. $c_1 < x'_1 \wedge y'_1$).

Proof. It is clear from (2.9) that (2.10) is increasing in c_i for every $i \neq 1$. Moreover, if $x'_1 x'_2 y'_1 y'_2 \neq 0$, then (2.10) is strictly increasing in c_2 when $c_1 < x'_1 \wedge y'_1$ since the latter implies $(x'_1 - c_1)(y'_1 - c_1) \neq 0$. Finally, regarding c_1 , observe that

$$\frac{d}{dc_1} \left(\frac{(x'_1 - c_1)(y'_1 - c_1)}{n' - (c_1 + \dots + c_k)} + c_1 \right) = \left(\frac{x'_1 - c_1}{n' - (c_1 + \dots + c_k)} - 1 \right) \left(\frac{y'_1 - c_1}{n' - (c_1 + \dots + c_k)} - 1 \right) \geq 0.$$

If $x'_1 x'_2 y'_1 y'_2 \neq 0$, then this inequality is strict when $c_2 < x'_2 \wedge y'_2$. This concludes the proof. \square

LEMMA 2.9. For every mating preference matrix P and every $\mathbf{t} \in \Phi$, we have

$$\mathbb{P}^{\mathbf{t}}(Q(t_1^*) \neq 0) \geq \min_{1 \leq i, j \leq k} p_{ij} > 0.$$

Proof. Consider the first animal that fires at the first round. If it forms a permanent pair at that time, then $Q(t_1^*) \neq 0$. The former event happens with probability at least $\min_{1 \leq i, j \leq k} p_{ij}$. \square

THEOREM 2.10. Assume that, for every $i, j \in \{1, \dots, k\}$, $p_{ij} = 1$ holds if and only if $i \neq j$. Then, we have the following results:

- (a) For every $i \in \{1, \dots, k\}$, $x_1, \dots, x_k, y_1, \dots, y_k \geq 0$ satisfying (1.1), and $\mathbf{t} \in \Phi$,

$$u_{ii}^*(x_1, \dots, x_k; y_1, \dots, y_k \mid \mathbf{t}) \leq \frac{x_i y_i}{n}.$$

In particular,

$$u_{ii}^*(x_1, \dots, x_k; y_1, \dots, y_k) \leq \frac{x_i y_i}{n}.$$

- (b) The inequalities in part (a) are strict whenever $x_i x_{i'} y_i y_{i'} \neq 0$ for some $i' \neq i$.

Proof. Since we can relabel the types, it suffices to prove the desired results for u_{11}^* . We start with part (a) and proceed by strong induction on $n \geq 1$. The case $n = 1$ is trivial. Indeed,

$$u_{11}^*(x_1, \dots, x_k; y_1, \dots, y_k \mid \mathbf{t}) = \begin{cases} 0 & \text{if } x_1 y_1 = 0, \text{ and} \\ 1 & \text{if } x_1 y_1 = 1. \end{cases}$$

For $n \geq 2$, let

$$(2.11) \quad \bar{u}_{11}^*(x_1, \dots, x_k; y_1, \dots, y_k) := \sup_{\mathbf{t} \in \Phi} u_{11}^*(x_1, \dots, x_k; y_1, \dots, y_k \mid \mathbf{t}).$$

With this notation, for every $\mathbf{t} \in \Phi$,

$$\begin{aligned}
(2.12) \quad & u_{11}^*(x_1, \dots, x_k; y_1, \dots, y_k \mid \mathbf{t}) \\
& \leq \sum_{M \in \mathcal{E} \setminus (\mathcal{E}' \cup \{0\})} \mathbb{P}^{\mathbf{t}}(Q(t_1^*) = M) \left[\frac{(x_1 - m_{1,\cdot})(y_1 - m_{\cdot,1})}{n - m_{tot}} + m_{11} \right] \\
& \quad + \sum_{M \in \mathcal{E}'} \mathbb{P}^{\mathbf{t}}(Q(t_1^*) = M) m_{11} + \mathbb{P}^{\mathbf{t}}(Q(t_1^*) = 0) \bar{u}_{11}^*(x_1, \dots, x_k; y_1, \dots, y_k) \\
& = \sum_{\substack{M \in \mathcal{E}: \\ M \neq 0}} \mathbb{P}^{\mathbf{t}}(Q(t_1^*) = M) [H(x_1 - m_{1,\cdot}, \dots, x_k - m_{k,\cdot}; y_1 - m_{\cdot,1}, \dots, y_k - m_{\cdot,k}) + m_{11}] \\
& \quad + \mathbb{P}^{\mathbf{t}}(Q(t_1^*) = 0) \bar{u}_{11}^*(x_1, \dots, x_k; y_1, \dots, y_k) \\
& = \sum_{M \in \mathcal{E}} \mathbb{P}^{\mathbf{t}}(Q(t_1^*) = M) [H(x_1 - m_{1,\cdot}, \dots, x_k - m_{k,\cdot}; y_1 - m_{\cdot,1}, \dots, y_k - m_{\cdot,k}) + m_{11}] \\
& \quad - \mathbb{P}^{\mathbf{t}}(Q(t_1^*) = 0) \frac{x_1 y_1}{n} + \mathbb{P}^{\mathbf{t}}(Q(t_1^*) = 0) \bar{u}_{11}^*(x_1, \dots, x_k; y_1, \dots, y_k) \\
(2.13) \quad & = \sum_{M \in \mathcal{E}} \sum_{M'} \mathbb{P}^{\mathbf{t}}(Q(t_1^*) = M, Q'(t_1^*) = M') \\
& \quad \times [H(x_1 - m_{1,\cdot}, \dots, x_k - m_{k,\cdot}; y_1 - m_{\cdot,1}, \dots, y_k - m_{\cdot,k}) + m_{11}] \\
& \quad - \mathbb{P}^{\mathbf{t}}(Q(t_1^*) = 0) \frac{x_1 y_1}{n} + \mathbb{P}^{\mathbf{t}}(Q(t_1^*) = 0) \bar{u}_{11}^*(x_1, \dots, x_k; y_1, \dots, y_k) \\
(2.14) \quad & \leq \sum_{M \in \mathcal{E}} \sum_{M'} \mathbb{P}^{\mathbf{t}}(Q(t_1^*) = M, Q'(t_1^*) = M') \\
& \quad \times [H(x_1 - m'_{1,\cdot}, \dots, x_k - m'_{k,\cdot}; y_1 - m'_{\cdot,1}, \dots, y_k - m'_{\cdot,k}) + m'_{11}] \\
& \quad - \mathbb{P}^{\mathbf{t}}(Q(t_1^*) = 0) \frac{x_1 y_1}{n} + \mathbb{P}^{\mathbf{t}}(Q(t_1^*) = 0) \bar{u}_{11}^*(x_1, \dots, x_k; y_1, \dots, y_k) \\
(2.15) \quad & = \sum_{M' \in \mathcal{E}} \mathbb{P}^{\mathbf{t}}(Q'(t_1^*) = M') [H(x_1 - m'_{1,\cdot}, \dots, x_k - m'_{k,\cdot}; y_1 - m'_{\cdot,1}, \dots, y_k - m'_{\cdot,k}) + m'_{11}] \\
& \quad - \mathbb{P}^{\mathbf{t}}(Q(t_1^*) = 0) \frac{x_1 y_1}{n} + \mathbb{P}^{\mathbf{t}}(Q(t_1^*) = 0) \bar{u}_{11}^*(x_1, \dots, x_k; y_1, \dots, y_k) \\
(2.16) \quad & = \frac{x_1 y_1}{n} - \mathbb{P}^{\mathbf{t}}(Q(t_1^*) = 0) \frac{x_1 y_1}{n} + \mathbb{P}^{\mathbf{t}}(Q(t_1^*) = 0) \bar{u}_{11}^*(x_1, \dots, x_k; y_1, \dots, y_k) \\
(2.17) \quad & = \mathbb{P}^{\mathbf{t}}(Q(t_1^*) \neq 0) \frac{x_1 y_1}{n} + \mathbb{P}^{\mathbf{t}}(Q(t_1^*) = 0) \bar{u}_{11}^*(x_1, \dots, x_k; y_1, \dots, y_k).
\end{aligned}$$

The inequality in (2.12) follows from the induction hypothesis since $n - m_{tot} \leq n - 1$ for $M \neq 0$. (Technically, one should first condition on $\mathcal{L}(t_1^*)$, and then use the induction hypothesis with the shifted firing times $(t_s(\zeta) - t_1^*)_{s \geq 1, \zeta \in S(t_1^*)}$ of the single animals.) $Q'(t_1^*)$ is the $k \times k$ matrix with ij th entry

$$Q'_{ij}(t_1^*) := \sum_{a=1}^n \sum_{b=1}^n \mathbb{I}_{\{(\varphi_a, \mu_b) \in (\Delta \mathcal{L})'(t_1^*), \Gamma(\varphi_a, \mu_b) = (i, j)\}}$$

which is the number of type- ij temporary pairs formed at the first firing round. Since there is definite mating upon encounter for mixed-type temporary pairs, the second sum in (2.13) is over all $M' \in \mathcal{E}$ such that $M' \geq M$ and $m'_{ij} = m_{ij}$ for every $i \neq j$. Lemma 2.8 with

$$x'_i = x_i - m_{i,\cdot}, \quad y'_j = y_j - m_{\cdot,j} \quad \text{and} \quad c_i = m'_{ii} - m_{ii}$$

gives (2.14). Note that M does not appear inside the square brackets in (2.14), so we can change the order of summation and obtain (2.15). Finally, the equality in (2.16) follows from

Corollary 2.5. Indeed, in the case of definite mating upon encounter for all temporary pairs, we have $Q'(t_1^*) = Q(t_1^*)$ and

$$\begin{aligned} u_{11}^*(x_1 - m'_{1,\cdot}, \dots, x_k - m'_{k,\cdot}; y_1 - m'_{\cdot,1}, \dots, y_k - m'_{\cdot,k} | \mathbf{t}) \\ = H(x_1 - m'_{1,\cdot}, \dots, x_k - m'_{k,\cdot}; y_1 - m'_{\cdot,1}, \dots, y_k - m'_{\cdot,k}) \end{aligned}$$

for every $M' \in \mathcal{E}$ and $\mathbf{t} \in \Phi$. The decomposition of $u_{11}^*(x_1, \dots, x_k; y_1, \dots, y_k | \mathbf{t}) = \frac{x_1 y_1}{n}$ via conditioning on $Q(t_1^*)$ gives the desired equality.

For every $\epsilon > 0$, there exists a $\mathbf{t}' \in \Phi$ such that

$$\bar{u}_{11}^*(x_1, \dots, x_k; y_1, \dots, y_k) \leq u_{11}^*(x_1, \dots, x_k; y_1, \dots, y_k | \mathbf{t}') + \epsilon,$$

and (2.17) gives

$$\mathbb{P}^{\mathbf{t}'}(Q(t_1^*) \neq 0) \bar{u}_{11}^*(x_1, \dots, x_k; y_1, \dots, y_k) \leq \mathbb{P}^{\mathbf{t}'}(Q(t_1^*) \neq 0) \frac{x_1 y_1}{n} + \epsilon.$$

By Lemma 2.9, we get

$$\bar{u}_{11}^*(x_1, \dots, x_k; y_1, \dots, y_k) \leq \frac{x_1 y_1}{n} + \frac{\epsilon}{\min\{p_{11}, \dots, p_{kk}\}}.$$

Since $\epsilon > 0$ is arbitrary, we deduce that

$$(2.18) \quad \bar{u}_{11}^*(x_1, \dots, x_k; y_1, \dots, y_k) \leq \frac{x_1 y_1}{n}.$$

This concludes the proof of part (a).

We prove part (b) by strong induction, too. The case $n = 1$ is vacuous since $x_1 x_i y_1 y_i = 0$ for every $i \neq 1$. For $n \geq 2$, note that if the supremum in (2.11) is not attained, then (2.18) implies the desired strict inequality. If the supremum in (2.11) is attained at some $\mathbf{t}' \in \Phi$, then we will assume WLOG that $x_1 x_2 y_1 y_2 \neq 0$ and consider two subcases:

- (i) If no type-1 or type-2 animal fires at the first round, then there exists an $M \in \mathcal{E}$ such that $M \neq 0$,

$$(2.19) \quad (x_1 - m_{1,\cdot})(x_2 - m_{2,\cdot})(y_1 - m_{\cdot,1})(y_2 - m_{\cdot,2}) \neq 0,$$

and $\mathbb{P}^{\mathbf{t}'}(Q(t_1^*) = M) > 0$. (Indeed, in the alternative representation of the SEM model, with positive probability, at least one type-1 (resp. type-2) female is paired up with a type-1 (resp. type-2) male. Since none of these animals fire at the first round, these pairs are discarded at the encounter stage.) Therefore, the inequality in (2.12) is strict by the induction hypothesis.

- (ii) If at least one type-1 or type-2 animal fires at the first round, then there exist $M, M' \in \mathcal{E}$ such that $m_{11} = m_{22} = 0$, $m'_{11} \vee m'_{22} \neq 0$, (2.19) holds, and

$$\mathbb{P}^{\mathbf{t}'}(Q(t_1^*) = M, Q'(t_1^*) = M') > 0.$$

(Indeed, with positive probability, at least one type-1 (resp. type-2) female is paired up with a type-1 (resp. type-2) male and each of these pairs is discarded at the encounter stage or the mating stage.) Therefore, the inequality in (2.14) is strict by Lemma 2.8.

In both of these subcases, we deduce that (2.18) holds with strict inequality. This concludes the proof of part (b). \square

REMARK 2.11. The key step in the proof of Theorem 2.10 is the inequality in (2.14) which follows from Lemma 2.8. This lemma says that, in the case of definite mating upon encounter, the expected number of type-11 pairs in the terminal pair-list would decrease if we were to interfere with the process at any time and discard some pure-type pairs. This is precisely how we are viewing what happens (with positive probability) at the first firing round.

REMARK 2.12. It is perhaps natural to look for an easier proof of Theorem 2.10 via a coupling with definite mating upon encounter. However, there are some obstacles to such approaches. For example, consider the following setup: $k = 2$, $n = 3$, $x_1 = y_1 = 2$, $x_2 = y_2 = 1$, $\gamma_1(\varphi_1) = \gamma_1(\varphi_2) = \gamma_2(\mu_1) = \gamma_2(\mu_2) = 1$, and $\gamma_1(\varphi_3) = \gamma_2(\mu_3) = 2$. Suppose the firing

times are ordered as $t_1(\varphi_1) < t_2(\varphi_1) < t_1(\mu_1) < t_1(\mu_2) < \dots$ and φ_1 samples μ_1 at $t_1^* = t_1(\varphi_1)$. In the case of definite mating upon encounter, if μ_2 samples φ_3 at $t_2^* = t_1(\mu_2)$, then $Q_{11}(T) = 1$. However, in the case of definite mating upon encounter for mixed-type temporary pairs only, if (i) the pair (φ_1, μ_1) is discarded, (ii) φ_1 samples μ_2 at $t_2^* = t_2(\varphi_1)$ and (iii) μ_1 samples φ_2 at $t_3^* = t_1(\mu_1)$, then $Q_{11}(T) = 2 > 1$, i.e., the inequality is in the reverse direction.

REMARK 2.13. Theorem 2.10 can be slightly generalized. For example, if we assume that $k \geq 3$, $p_{11} < 1$, $p_{22} \leq 1$, $p_{ii} = 1$ for every $i \in \{3, \dots, k\}$, and $p_{ij} = 1$ for every $i, j \in \{1, \dots, k\}$ such that $i \neq j$, then

$$(2.20) \quad u_{11}^*(x_1, \dots, x_k; y_1, \dots, y_k | \mathbf{t}) \leq \frac{x_1 y_1}{n}$$

still holds, and this inequality is strict when $p_{22} < 1$ and $x_1 x_2 y_1 y_2 \neq 0$. However, the inequality in (2.20) is not necessarily strict when $p_{22} = 1$ and $x_1 x_2 y_1 y_2 \neq 0$. As a counterexample, assume that all males fire at the first round. Since $p_{ij} < 1$ if and only if $i = j = 1$, there are only type-1 females and males in the singles' pool $S(t_1^*)$, and these animals have no option other than eventually forming type-11 permanent pairs. Therefore, the distribution of $Q(T)$ under $\mathbb{P}^{\mathbf{t}}$ is identical to what it was in the case of definite mating upon encounter for all temporary pairs. In particular, we have equality in (2.20).

3. POISSON FIRING TIMES

3.1. The infinitesimal generator. In this section, we will make the following assumptions regarding the firing times of the animals.

(Poi1) $\{N(\zeta)\}_{\zeta \in Z}$ are mutually independent Poisson processes on $[0, \infty)$.

(Poi2) The intensity of $N(\zeta)$ is equal to some $\alpha_i \geq 0$ (resp. $\beta_j \geq 0$) for all type- i females (resp. type- j males).

(Poi3) $\alpha_i + \beta_j > 0$ for every $i, j \in \{1, \dots, k\}$.

Note that (Poi1)–(Poi3) are stronger than (Gen1)–(Gen4) in Section 2. Indeed, (Gen3) is automatically satisfied for Poisson processes, and (Poi3) implies (Gen4).

In this case, since $\{N(\zeta)\}_{\zeta \in Z}$ are memoryless, it is clear that the pair-type process $Q(\cdot)$ is a continuous-time pure jump Markov chain under \mathbb{P} , with state space \mathcal{E} defined in (2.1). Moreover, the jumps of $Q_{ij}(\cdot)$ are of size 1 because almost surely one animal fires at each firing round. Therefore, $Q(\cdot)$ may be regarded as a multidimensional pure birth process, see [19].

PROPOSITION 3.1. *Assume (Poi1)–(Poi3). Take a population with x_1, \dots, x_k females and y_1, \dots, y_k males of types $1, \dots, k$, respectively, such that (1.1) holds. The infinitesimal generator of the continuous-time Markov chain $Q(\cdot)$ has the following formula: for every $M, M' \in \mathcal{E}$,*

$$(3.1) \quad \begin{aligned} \rho(M, M') &:= \lim_{\Delta t \rightarrow 0^+} \frac{1}{\Delta t} \mathbb{P}(Q(t + \Delta t) = M' | Q(t) = M) \\ &= \begin{cases} \frac{\pi_{ij}(x_i - m_{i,\cdot})(y_j - m_{\cdot,j})}{n - m_{tot}} & \text{if } M' = M + I^{ij}; \text{ and} \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Here, $I^{ij} \in \mathcal{M}^{k \times k}(\mathbb{N} \cup \{0\})$ denotes the $k \times k$ matrix whose entries are zero except its ij th entry which is 1, and

$$(3.2) \quad \pi_{ij} := p_{ij}(\alpha_i + \beta_j).$$

Proof. Conditioned on $Q(t) = M$, there are $x_i - m_{i,\cdot}$ single type- i females at time t and each of them fires independently with probability $\alpha_i \Delta t + o(\Delta t)$ in $[t, t + \Delta t]$. Conditioned on the latter event, each single type- i female samples a single type- j male with probability $\frac{y_j - m_{\cdot,j}}{n - m_{tot}}$. This is one of the two ways of forming a temporary type- ij pair. The other way is obtained by switching the roles of females and males. Combining these two ways, we see that the conditional probability of forming a temporary type- ij pair in $[t, t + \Delta t]$ given $Q(t) = M$ is

$$\frac{(\alpha_i \Delta t + \beta_j \Delta t + o(\Delta t))(x_i - m_{i,\cdot})(y_j - m_{\cdot,j})}{n - m_{tot}}.$$

Finally, each temporary type- ij pair becomes a permanent pair with probability p_{ij} , and this implies (3.1). \square

REMARK 3.2. It follows readily from (3.1) that the distribution of the pair-type process $Q(\cdot)$ depends on $P = (p_{ij})$, $\alpha_1, \dots, \alpha_k$ and β_1, \dots, β_k only through the $k \times k$ matrix

$$\Pi := (\pi_{ij})$$

whose entries are defined in (3.2). In other words, we have the freedom to change these parameters as long as Π stays the same. Therefore, under the assumption of Poisson firing times, the EM law of the species can be identified with Π .

Since the pair-type process $Q(\cdot)$ is a time-homogeneous Markov chain with infinitesimal generator ρ as in (3.1), the Kolmogorov backward equation for (1.3) is

$$\begin{aligned} (3.3) \quad & \frac{\partial}{\partial t} u_{ij}(t; x_1, \dots, x_k; y_1, \dots, y_k) \\ &= \sum_{i'=1}^k \sum_{j'=1}^k \frac{\pi_{i'j'} x_{i'} y_{j'}}{n} [u_{ij}(t; x_1 - \delta_{1i'}, \dots, x_k - \delta_{ki'}, y_1 - \delta_{1j'}, \dots, y_k - \delta_{kj'}) + \delta_{ii'} \delta_{jj'}] \\ & \quad - \frac{z}{n} u_{ij}(t; x_1, \dots, x_k; y_1, \dots, y_k). \end{aligned}$$

Here, $\delta_{..}$ denotes the Kronecker delta function, and

$$z := \sum_{i=1}^k \sum_{j=1}^k \pi_{ij} x_i y_j.$$

Note that (3.3) is a closed system of recursive ODEs.

Similarly, using the strong Markov property, we can decompose (1.4) w.r.t. the possible values of $Q(t_1^*(\tau))$ and thereby obtain the following recursion:

$$\begin{aligned} (3.4) \quad & u_{ij}^*(x_1, \dots, x_k; y_1, \dots, y_k) \\ &= \sum_{i'=1}^k \sum_{j'=1}^k \frac{\pi_{i'j'} x_{i'} y_{j'}}{z} [u_{ij}^*(x_1 - \delta_{1i'}, \dots, x_k - \delta_{ki'}, y_1 - \delta_{1j'}, \dots, y_k - \delta_{kj'}) + \delta_{ii'} \delta_{jj'}]. \end{aligned}$$

This was previously given by Mosteller [20].

The recursive equations (3.3) and (3.4) can be solved numerically, especially when n is not too large. However, it is not clear if they can be solved analytically. In the next subsection, we will provide explicit solutions to these equations when the EM law Π satisfies the so-called Poisson fine balance condition.

3.2. Poisson fine balance and panmixia.

DEFINITION 3.3. *The EM law Π is said to satisfy the Poisson fine balance condition if*

$$(3.5) \quad \pi_{ij} + \pi_{i'j'} = \pi_{ij'} + \pi_{i'j}$$

for every $i, j, i', j' \in \{1, \dots, k\}$.

The motivation behind this definition is the following observation.

PROPOSITION 3.4. *Assume (Poi1)–(Poi3). If the species is panmictic, then the EM law Π satisfies the Poisson fine balance condition.*

Proof. Assume that the species is panmictic. Observe that (3.5) is trivial if $i = i'$ or $j = j'$. Take a population with $n = 2$ and $x_1 = x_2 = y_1 = y_2 = 1$. Use recursion (3.4) to write

$$\begin{aligned} \frac{1}{2} &= \frac{x_1 y_1}{n} = u_{11}^*(x_1, \dots, x_k; y_1, \dots, y_k) \\ &= \sum_{i=1}^k \sum_{j=1}^k \frac{\pi_{ij} x_i y_j}{z} [u_{11}^*(x_1 - \delta_{1i}, \dots, x_k - \delta_{ki}; y_1 - \delta_{1j}, \dots, y_k - \delta_{kj}) + \delta_{1i} \delta_{1j}] \\ &= \frac{1}{z} (\pi_{11}[0+1] + \pi_{12}[0+0] + \pi_{21}[0+0] + \pi_{22}[1+0]) \\ &= \frac{\pi_{11} + \pi_{22}}{\pi_{11} + \pi_{12} + \pi_{21} + \pi_{22}}. \end{aligned}$$

Therefore, $\pi_{11} + \pi_{22} = \pi_{12} + \pi_{21}$. This is (3.5) with $i = j = 1$ and $i' = j' = 2$. By relabeling the types, the same argument works for any $i, j, i', j' \in \{1, \dots, k\}$ such that $i \neq i'$ and $j \neq j'$, and we are done. \square

It is natural to ask if the converse is true, i.e., whether the species is panmictic whenever the EM law satisfies the Poisson fine balance condition. This turns out to be true and can be verified by induction and (3.4). However, we will prove it by taking a much more conceptual approach. First, we need a lemma.

LEMMA 3.5. *The EM law Π satisfies the Poisson fine balance condition if and only if there exist $\bar{\alpha}_1, \dots, \bar{\alpha}_k, \bar{\beta}_1, \dots, \bar{\beta}_k \geq 0$ such that*

$$(3.6) \quad \pi_{ij} = \bar{\alpha}_i + \bar{\beta}_j$$

for every $i, j \in \{1, \dots, k\}$.

Proof. Assume that (3.6) indeed holds with some $\bar{\alpha}_1, \dots, \bar{\alpha}_k, \bar{\beta}_1, \dots, \bar{\beta}_k \geq 0$. Then,

$$\pi_{ij} + \pi_{i'j'} = (\bar{\alpha}_i + \bar{\beta}_j) + (\bar{\alpha}_{i'} + \bar{\beta}_{j'}) = (\bar{\alpha}_i + \bar{\beta}_{j'}) + (\bar{\alpha}_{i'} + \bar{\beta}_j) = \pi_{ij'} + \pi_{i'j}$$

for every $i, j, i', j' \in \{1, \dots, k\}$, i.e., the Poisson fine balance condition is satisfied.

Conversely, if the Poisson fine balance condition holds, then assume WLOG (i.e., up to relabeling the male types) that $\pi_{11} = \min\{\pi_{11}, \dots, \pi_{1k}\}$. Define

$$\bar{\alpha}_i := \pi_{i1} > 0 \quad \text{and} \quad \bar{\beta}_j := \pi_{1j} - \pi_{11} \geq 0.$$

With this notation, we have

$$\pi_{ij} = \pi_{i1} + \pi_{1j} - \pi_{11} = \bar{\alpha}_i + \bar{\beta}_j,$$

where the first equality follows from (3.5). This concludes the proof. \square

THEOREM 3.6. *Assume (Poi1)–(Poi3). Take a population with x_1, \dots, x_k females and y_1, \dots, y_k males of types $1, \dots, k$, respectively, such that (1.1) holds. If the EM law Π satisfies the Poisson fine balance condition, then*

$$(3.7) \quad \mathbb{P}(Q(t) = M) = \frac{\left(\prod_i x_i!\right) \left(\prod_j y_j!\right)}{n! \left(\prod_{i,j} m_{ij}!\right)} \sum_{\substack{M' \in \mathcal{E}' \\ M' \geq M}} \prod_{i,j} \frac{(1 - e^{-\pi_{ij}t})^{m_{ij}} (e^{-\pi_{ij}t})^{m'_{ij} - m_{ij}}}{(m'_{ij} - m_{ij})!}$$

for every $t \geq 0$ and $M \in \mathcal{E}$. Similarly,

$$(3.8) \quad \mathbb{P}(Q(T) = M) = \frac{\left(\prod_i x_i!\right) \left(\prod_j y_j!\right)}{n! \left(\prod_{i,j} m_{ij}!\right)},$$

$$(3.9) \quad u_{ij}^*(x_1, \dots, x_k; y_1, \dots, y_k) = \frac{x_i y_j}{n}, \quad \text{and}$$

$$(3.10) \quad u_{ij}(t; x_1, \dots, x_k; y_1, \dots, y_k) = \frac{x_i y_j}{n} (1 - e^{-\pi_{ij}t}).$$

Proof. If the EM law satisfies the Poisson fine balance condition, then Lemma 3.5 implies that (3.6) holds with some $\bar{\alpha}_1, \dots, \bar{\alpha}_k, \bar{\beta}_1, \dots, \bar{\beta}_k \geq 0$. In the light of Remark 3.2, as far as the distribution of the pair-type process $Q(\cdot)$ is concerned, we can change (if necessary) $P = (p_{ij})$, $\alpha_1, \dots, \alpha_k$ and β_1, \dots, β_k , and assume that

$$p_{ij} = 1, \quad \alpha_i = \bar{\alpha}_i \quad \text{and} \quad \beta_j = \bar{\beta}_j$$

for every $i, j \in \{1, \dots, k\}$. But then, we have definite mating upon encounter and Theorem 2.6 is applicable. Since $\tau_1(\zeta)$ is exponentially distributed with rate $\bar{\alpha}_i$ (resp. $\bar{\beta}_j$) if ζ is a type- i female (resp. type- j male), we have

$$\lambda_{ij}(t) = 1 - (1 - F_i(t))(1 - G_j(t)) = 1 - e^{-\bar{\alpha}_i t} e^{-\bar{\beta}_j t} = 1 - e^{-\pi_{ij} t}$$

and this implies (3.7). Similarly, Corollaries 2.4, 2.5 and 2.7 give (3.8), (3.9) and (3.10), respectively. \square

COROLLARY 3.7. *Assume (Poi1)–(Poi3). The species is panmictic if and only if the EM law Π satisfies the Poisson fine balance condition.*

Proof. This is immediate from Proposition 3.4 and Theorem 3.6. \square

3.3. Full characterization of panmixia, homogamy and heterogamy in the 2×2 case.

LEMMA 3.8. *For $k = 2$, the EM law Π satisfies $\pi_{11} + \pi_{22} < \pi_{12} + \pi_{21}$ if and only if there exist $\bar{\alpha}_1, \bar{\alpha}_2, \bar{\beta}_1, \bar{\beta}_2 \geq 0$ such that*

$$(3.11) \quad \pi_{11} < \bar{\alpha}_1 + \bar{\beta}_1, \quad \pi_{12} = \bar{\alpha}_1 + \bar{\beta}_2, \quad \pi_{21} = \bar{\alpha}_2 + \bar{\beta}_1 \quad \text{and} \quad \pi_{22} < \bar{\alpha}_2 + \bar{\beta}_2.$$

Proof. Assume that (3.11) indeed holds with some $\bar{\alpha}_1, \bar{\alpha}_2, \bar{\beta}_1, \bar{\beta}_2 \geq 0$. Then,

$$\pi_{11} + \pi_{22} < (\bar{\alpha}_1 + \bar{\beta}_1) + (\bar{\alpha}_2 + \bar{\beta}_2) = (\bar{\alpha}_1 + \bar{\beta}_2) + (\bar{\alpha}_2 + \bar{\beta}_1) = \pi_{12} + \pi_{21}.$$

Conversely, if $\pi_{11} + \pi_{22} < \pi_{12} + \pi_{21}$, then pick $\bar{\Pi} = (\bar{\pi}_{ij})$ such that

$$\pi_{11} < \bar{\pi}_{11}, \quad \pi_{12} = \bar{\pi}_{12}, \quad \pi_{21} = \bar{\pi}_{21}, \quad \pi_{22} < \bar{\pi}_{22} \quad \text{and} \quad \bar{\pi}_{11} + \bar{\pi}_{22} = \bar{\pi}_{12} + \bar{\pi}_{21}.$$

The desired result follows from Lemma 3.5 applied to $\bar{\Pi}$. \square

THEOREM 3.9. *Assume (Poi1)–(Poi3). For $k = 2$, the species is*

- (i) *heterogamous if* $\pi_{11} + \pi_{22} < \pi_{12} + \pi_{21}$,
- (ii) *panmictic if* $\pi_{11} + \pi_{22} = \pi_{12} + \pi_{21}$, *and*
- (iii) *homogamous if* $\pi_{11} + \pi_{22} > \pi_{12} + \pi_{21}$.

Proof. Let us prove part (i). If $\pi_{11} + \pi_{22} < \pi_{12} + \pi_{21}$, then Lemma 3.8 implies that (3.11) holds with some $\bar{\alpha}_1, \bar{\alpha}_2, \bar{\beta}_1, \bar{\beta}_2 \geq 0$. In the light of Remark 3.2, as far as the distribution of the pair-type process $Q(\cdot)$ is concerned, we can change (if necessary) $P = (p_{ij})$, α_1, α_2 and β_1, β_2 , and assume that

$$p_{11} < 1, \quad p_{12} = 1, \quad p_{21} = 1, \quad p_{22} < 1, \quad \alpha_i = \bar{\alpha}_i \quad \text{and} \quad \beta_j = \bar{\beta}_j$$

for every $i, j \in \{1, 2\}$. But then, we have definite mating upon encounter for mixed-type temporary pairs only, and Theorem 2.10 is applicable. This concludes the proof of part (i). We have already shown part (ii) in Corollary 3.7. Finally, part (iii) follows from part (i) by relabeling the male types. \square

4. BERNOULLI FIRING TIMES

4.1. The transition kernel. In this section, we will make the following assumptions.

- (Ber1) $\{N(\zeta)\}_{\zeta \in Z}$ are mutually independent Bernoulli processes on \mathbb{N} .
- (Ber2) The success probability of the Bernoulli trials of $N(\zeta)$ is equal to some $\alpha_i \in [0, 1]$ (resp. $\beta_j \in [0, 1]$) for all type- i females (resp. type- j males).
- (Ber3) $\alpha_i + \beta_j > 0$ for every $i, j \in \{1, \dots, k\}$.

Note that (Ber1)–(Ber3) are stronger than (Gen1)–(Gen4) in Section 2. Indeed, (Gen3) is automatically satisfied as the Bernoulli processes are defined on \mathbb{N} , and (Ber3) implies (Gen4).

Similar to the Poisson case, since $\{N(\zeta)\}_{\zeta \in Z}$ are memoryless, it is clear that the pair-type process $Q(\cdot)$ is a discrete-time Markov chain under \mathbb{P} , with state space \mathcal{E} defined in (2.1). However, the jumps of $Q_{ij}(\cdot)$ are not necessarily of size 1 because more than one animal can fire at each firing round.

PROPOSITION 4.1. *Assume (Ber1)–(Ber3). Take a population with x_1, \dots, x_k females and y_1, \dots, y_k males of types $1, \dots, k$, respectively, such that (1.1) holds. The transition kernel of the discrete-time Markov chain $Q(\cdot)$ has the following formula: for every $M, M' \in \mathcal{E}$,*

$$(4.1) \quad \begin{aligned} \rho(M, M') &:= \mathbb{P}(Q(t+1) = M' \mid Q(t) = M) \\ &= \frac{\left(\prod_i (x_i - m_{i,\cdot})!\right) \left(\prod_j (y_j - m_{\cdot,j})!\right)}{(n - m_{tot})! \left(\prod_{i,j} (m'_{ij} - m_{ij})!\right)} \sum_{\substack{M'' \in \mathcal{E}: \\ M'' \geq M'}} \prod_{i,j} \frac{(\pi_{ij})^{m'_{ij} - m_{ij}} (1 - \pi_{ij})^{m''_{ij} - m'_{ij}}}{(m''_{ij} - m'_{ij})!} \end{aligned}$$

if $M' \geq M$; and 0 otherwise. Here,

$$(4.2) \quad \pi_{ij} := p_{ij}(\alpha_i + \beta_j - \alpha_i \beta_j) \in (0, 1].$$

Proof. By the Markov property, conditioning on $Q(t) = M$ is equivalent to taking $t = 0$ and assuming that there are initially $x_i - m_{i,\cdot}$ single type- i females and $y_j - m_{\cdot,j}$ single type- j males for every $i, j \in \{1, \dots, k\}$. We would like to find the distribution of $Q(1)$ under these assumptions. To this end, let $Q'(\cdot)$ be the pair-type process in the case of definite mating upon encounter. As we have seen in the proof of Theorem 2.6,

$$Q'_{ij}(1) \mid Q'_{ij}(T) \sim B(Q'_{ij}(T), \lambda_{ij}(1)),$$

where B denotes the binomial distribution and

$$\lambda_{ij}(1) = F_i(1) + G_j(1) - F_i(1)G_j(1) = \alpha_i + \beta_j - \alpha_i \beta_j.$$

Now, recall that

$$Q_{ij}(1) \mid Q'_{ij}(1) \sim B(Q'_{ij}(1), p_{ij}).$$

Therefore, by thinning,

$$Q_{ij}(1) \mid Q'_{ij}(T) \sim B(Q'_{ij}(T), \pi_{ij}),$$

with $\pi_{ij} = p_{ij}\lambda_{ij}(1)$ as in (4.2), and (4.1) is obtained by modifying the proof of Theorem 2.6 accordingly. \square

REMARK 4.2. It is evident from (4.1) that the distribution of the pair-type process $Q(\cdot)$ depends on $P = (p_{ij})$, $\alpha_1, \dots, \alpha_k$ and β_1, \dots, β_k only through the $k \times k$ matrix $\Pi := (\pi_{ij})$ whose entries are defined in (4.2). In other words, we have the freedom to change these parameters as long as Π stays the same. Therefore, under the assumption of Bernoulli firing times, the EM law of the species can be identified with Π .

Since the pair-type process $Q(\cdot)$ is a time-homogeneous Markov chain with transition kernel ρ as in (4.1), the Kolmogorov backward equation for (1.3) is

$$(4.3) \quad \begin{aligned} &u_{ij}(t+1; x_1, \dots, x_k; y_1, \dots, y_k) \\ &= \sum_{M \in \mathcal{E}} \rho(0, M) [u_{ij}(t; x_1 - m_{1,\cdot}, \dots, x_k - m_{k,\cdot}; y_1 - m_{\cdot,1}, \dots, y_k - m_{\cdot,k}) + m_{ij}]. \end{aligned}$$

Note that (4.3) is a closed system of recursive difference equations.

Similarly, using the strong Markov property, we can decompose (1.4) w.r.t. the possible values of $Q(t_1^*(\tau))$ and thereby obtain the following recursion:

$$(4.4) \quad u_{ij}^*(x_1, \dots, x_k; y_1, \dots, y_k) = \sum_{\substack{M \in \mathcal{E}: \\ M \neq 0}} \frac{\rho(0, M)}{1 - \rho(0, 0)} [u_{ij}^*(x_1 - m_{1,\cdot}, \dots, x_k - m_{k,\cdot}; y_1 - m_{\cdot,1}, \dots, y_k - m_{\cdot,k}) + m_{ij}].$$

The recursive equations (4.3) and (4.4) can be solved numerically, especially when n is not too large. However, it is not clear if they can be solved analytically.

4.2. Bernoulli fine balance and panmixia.

DEFINITION 4.3. *The EM law Π is said to satisfy the Bernoulli fine balance condition if*

$$(4.5) \quad (1 - \pi_{ij})(1 - \pi_{i'j'}) = (1 - \pi_{ij'})(1 - \pi_{i'j})$$

for every $i, j, i', j' \in \{1, \dots, k\}$.

PROPOSITION 4.4. *Assume (Ber1)–(Ber3). If the species is panmictic, then the EM law Π satisfies the Bernoulli fine balance condition.*

Proof. Assume that the species is panmictic. Observe that (4.5) is trivial if $i = i'$ or $j = j'$. Take a population with $n = 2$ and $x_1 = x_2 = y_1 = y_2 = 1$. Use recursion (4.4) to write

$$\begin{aligned} \frac{1}{2} &= \frac{x_1 y_1}{n} = u_{11}^*(x_1, \dots, x_k; y_1, \dots, y_k) \\ &= \sum_{\substack{M \in \mathcal{E}: \\ M \neq 0}} \frac{\rho(0, M)}{1 - \rho(0, 0)} [u_{11}^*(x_1 - m_{1,\cdot}, \dots, x_k - m_{k,\cdot}; y_1 - m_{\cdot,1}, \dots, y_k - m_{\cdot,k}) + m_{11}] \\ &= \frac{\frac{1}{2}(\pi_{11}(1 - \pi_{22})[0 + 1] + (1 - \pi_{11})\pi_{22}[1 + 0] + \pi_{11}\pi_{22}[0 + 1])}{\frac{1}{2}(\pi_{11}(1 - \pi_{22}) + (1 - \pi_{11})\pi_{22} + \pi_{11}\pi_{22}) + \frac{1}{2}(\pi_{12}(1 - \pi_{21}) + (1 - \pi_{12})\pi_{21} + \pi_{12}\pi_{21})} \\ &= \frac{\pi_{11} + \pi_{22} - \pi_{11}\pi_{22}}{(\pi_{11} + \pi_{22} - \pi_{11}\pi_{22}) + (\pi_{12} + \pi_{21} - \pi_{12}\pi_{21})}. \end{aligned}$$

This follows from computing $\rho(0, M)$ for every

$$M \in \mathcal{E} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \right\}$$

and observing that only the first three of these matrices contribute to the sum above. Therefore, $\pi_{11} + \pi_{22} - \pi_{11}\pi_{22} = \pi_{12} + \pi_{21} - \pi_{12}\pi_{21}$. This is (4.5) with $i = j = 1$ and $i' = j' = 2$. By relabeling the types, the same argument works for any $i, j, i', j' \in \{1, \dots, k\}$ such that $i \neq i'$ and $j \neq j'$, and we are done. \square

LEMMA 4.5. *The EM law Π satisfies the Bernoulli fine balance condition if and only if there exist $\bar{\alpha}_1, \dots, \bar{\alpha}_k, \bar{\beta}_1, \dots, \bar{\beta}_k \in [0, 1]$ such that*

$$(4.6) \quad 1 - \pi_{ij} = (1 - \bar{\alpha}_i)(1 - \bar{\beta}_j)$$

for every $i, j \in \{1, \dots, k\}$.

Proof. Assume that (4.6) indeed holds with some $\bar{\alpha}_1, \dots, \bar{\alpha}_k, \bar{\beta}_1, \dots, \bar{\beta}_k \in [0, 1]$. Then,

$$\begin{aligned} (1 - \pi_{ij})(1 - \pi_{i'j'}) &= (1 - \bar{\alpha}_i)(1 - \bar{\beta}_j)(1 - \bar{\alpha}_{i'})(1 - \bar{\beta}_{j'}) \\ &= (1 - \bar{\alpha}_i)(1 - \bar{\beta}_{j'})(1 - \bar{\alpha}_{i'})(1 - \bar{\beta}_j) = (1 - \pi_{ij'})(1 - \pi_{i'j}) \end{aligned}$$

for every $i, j, i', j' \in \{1, \dots, k\}$, i.e., the Bernoulli fine balance condition is satisfied.

Conversely, assume that the Bernoulli fine balance condition holds. If $\pi_{ij} = 1$ for every $i, j \in \{1, \dots, k\}$, then (4.6) holds with $\bar{\alpha}_i = \bar{\beta}_j = 1$. Otherwise, assume WLOG (i.e., up to relabeling the female and male types) that $\pi_{11} = \min\{\pi_{11}, \dots, \pi_{1k}\} < 1$. Define

$$\bar{\alpha}_i := \pi_{i1} \in (0, 1] \quad \text{and} \quad \bar{\beta}_j := 1 - \frac{1 - \pi_{1j}}{1 - \pi_{11}} \in [0, 1].$$

With this notation, we have

$$1 - \pi_{ij} = \frac{(1 - \pi_{i1})(1 - \pi_{1j})}{1 - \pi_{11}} = (1 - \bar{\alpha}_i)(1 - \bar{\beta}_j),$$

where the first equality follows from (4.5). This concludes the proof. \square

THEOREM 4.6. *Assume (Ber1)–(Ber3). Take a population with x_1, \dots, x_k females and y_1, \dots, y_k males of types $1, \dots, k$, respectively, such that (1.1) holds. If the EM law Π satisfies the Bernoulli fine balance condition, then*

$$(4.7) \quad \mathbb{P}(Q(t) = M) = \frac{\left(\prod_i x_i!\right) \left(\prod_j y_j!\right)}{n! \left(\prod_{i,j} m_{ij}!\right)} \sum_{\substack{M' \in \mathcal{E}': \\ M' \geq M}} \prod_{i,j} \frac{(1 - (1 - \pi_{ij})^t)^{m_{ij}} ((1 - \pi_{ij})^t)^{m'_{ij} - m_{ij}}}{(m'_{ij} - m_{ij})!}$$

for every $t \in \mathbb{N} \cup \{0\}$ and $M \in \mathcal{E}$. Similarly,

$$(4.8) \quad \mathbb{P}(Q(T) = M) = \frac{\left(\prod_i x_i!\right) \left(\prod_j y_j!\right)}{n! \left(\prod_{i,j} m_{ij}!\right)},$$

$$(4.9) \quad u_{ij}^*(x_1, \dots, x_k; y_1, \dots, y_k) = \frac{x_i y_j}{n}, \quad \text{and}$$

$$(4.10) \quad u_{ij}(t; x_1, \dots, x_k; y_1, \dots, y_k) = \frac{x_i y_j}{n} (1 - (1 - \pi_{ij})^t).$$

Proof. If the EM law satisfies the Bernoulli fine balance condition, then Lemma 4.5 implies that (4.6) holds with some $\bar{\alpha}_1, \dots, \bar{\alpha}_k, \bar{\beta}_1, \dots, \bar{\beta}_k \in [0, 1]$, which is equivalent to

$$\pi_{ij} = \bar{\alpha}_i + \bar{\beta}_j - \bar{\alpha}_i \bar{\beta}_j.$$

In the light of Remark 4.2, as far as the distribution of the pair-type process $Q(\cdot)$ is concerned, we can change (if necessary) $P = (p_{ij})$, $\alpha_1, \dots, \alpha_k$ and β_1, \dots, β_k , and assume that

$$p_{ij} = 1, \quad \alpha_i = \bar{\alpha}_i \quad \text{and} \quad \beta_j = \bar{\beta}_j$$

for every $i, j \in \{1, \dots, k\}$. But then, we have definite mating upon encounter and Theorem 2.6 is applicable. Since $\tau_1(\zeta)$ is geometrically distributed with success probability $\bar{\alpha}_i$ (resp. $\bar{\beta}_j$) if ζ is a type- i female (resp. type- j male), we have

$$\lambda_{ij}(t) = 1 - (1 - F_i(t))(1 - G_j(t)) = 1 - (1 - \bar{\alpha}_i)^t (1 - \bar{\beta}_j)^t = 1 - (1 - \pi_{ij})^t$$

and this implies (4.7). Similarly, Corollaries 2.4, 2.5 and 2.7 give (4.8), (4.9) and (4.10), respectively. \square

COROLLARY 4.7. *Assume (Ber1)–(Ber3). The species is panmictic if and only if the EM law Π satisfies the Bernoulli fine balance condition.*

Proof. This is immediate from Proposition 4.4 and Theorem 4.6. \square

4.3. Full characterization of panmixia, homogamy and heterogamy in the 2×2 case.

LEMMA 4.8. *For $k = 2$, the EM law Π satisfies $(1 - \pi_{11})(1 - \pi_{22}) > (1 - \pi_{12})(1 - \pi_{21})$ if and only if there exist $\bar{\alpha}_1, \bar{\alpha}_2, \bar{\beta}_1, \bar{\beta}_2 \in [0, 1]$ such that*

$$(4.11) \quad \begin{aligned} 1 - \pi_{11} &> (1 - \bar{\alpha}_1)(1 - \bar{\beta}_1), & 1 - \pi_{12} &= (1 - \bar{\alpha}_1)(1 - \bar{\beta}_2), \\ 1 - \pi_{21} &= (1 - \bar{\alpha}_2)(1 - \bar{\beta}_1) & \text{and} & \quad 1 - \pi_{22} > (1 - \bar{\alpha}_2)(1 - \bar{\beta}_2). \end{aligned}$$

Proof. Assume that (4.11) indeed holds with some $\bar{\alpha}_1, \bar{\alpha}_2, \bar{\beta}_1, \bar{\beta}_2 \in [0, 1]$. Then,

$$\begin{aligned} (1 - \pi_{11})(1 - \pi_{22}) &> (1 - \bar{\alpha}_1)(1 - \bar{\beta}_1)(1 - \bar{\alpha}_2)(1 - \bar{\beta}_2) \\ &= (1 - \bar{\alpha}_1)(1 - \bar{\beta}_2)(1 - \bar{\alpha}_2)(1 - \bar{\beta}_1) = (1 - \pi_{12})(1 - \pi_{21}). \end{aligned}$$

Conversely, if $(1 - \pi_{11})(1 - \pi_{22}) > (1 - \pi_{12})(1 - \pi_{21})$, then pick $\bar{\Pi} = (\bar{\pi}_{ij})$ such that

$$\pi_{11} < \bar{\pi}_{11}, \quad \pi_{12} = \bar{\pi}_{12}, \quad \pi_{21} = \bar{\pi}_{21}, \quad \pi_{22} < \bar{\pi}_{22} \quad \text{and} \quad (1 - \bar{\pi}_{11})(1 - \bar{\pi}_{22}) = (1 - \bar{\pi}_{12})(1 - \bar{\pi}_{21}).$$

The desired result follows from Lemma 4.5 applied to $\bar{\Pi}$. \square

THEOREM 4.9. *Assume (Ber1)–(Ber3). For $k = 2$, the species is*

- (i) *heterogamous if* $(1 - \pi_{11})(1 - \pi_{22}) > (1 - \pi_{12})(1 - \pi_{21})$,
- (ii) *panmictic if* $(1 - \pi_{11})(1 - \pi_{22}) = (1 - \pi_{12})(1 - \pi_{21})$, and
- (iii) *homogamous if* $(1 - \pi_{11})(1 - \pi_{22}) < (1 - \pi_{12})(1 - \pi_{21})$.

Proof. Let us prove part (i). If $(1 - \pi_{11})(1 - \pi_{22}) > (1 - \pi_{12})(1 - \pi_{21})$, then Lemma 4.8 implies that (4.11) holds with some $\bar{\alpha}_1, \bar{\alpha}_2, \bar{\beta}_1, \bar{\beta}_2 \in [0, 1]$. In the light of Remark 4.2, as far as the distribution of the pair-type process $Q(\cdot)$ is concerned, we can change (if necessary) $P = (p_{ij})$, α_1, α_2 and β_1, β_2 , and assume that

$$p_{11} < 1, \quad p_{12} = 1, \quad p_{21} = 1, \quad p_{22} < 1, \quad \alpha_i = \bar{\alpha}_i \quad \text{and} \quad \beta_j = \bar{\beta}_j$$

for every $i, j \in \{1, 2\}$. But then, we have definite mating upon encounter for mixed-type temporary pairs only, and Theorem 2.10 is applicable. This concludes the proof of part (i). We have already shown part (ii) in Corollary 4.7. Finally, part (iii) follows from part (i) by relabeling the male types. \square

5. SOME CONCLUDING REMARKS, OBSERVATIONS, AND OPEN PROBLEMS

Main features of the model. Gimelfarb [13, p. 873] says: “Comparing first the applicability of the two encounter-mating models to real populations, it would seem that model 1 (individual encounters) is better suited for populations in which individuals are dispersed and search for potential mates individually, whereas model 2 (mass encounters) is better for populations in which individuals aggregate in search for potential mates, for example, mating swarms or leks. It is also possible that in some populations more than one encounter may take place at a time, and yet not all of the individuals available for mating participate in encounters simultaneously; thus, neither of the two models will describe such a population correctly.” The SEM model unifies and generalizes the individual and mass EM models, and thereby removes the aforesaid limitation regarding their scope. Moreover, by tuning the parameters $\alpha_1, \dots, \alpha_k$ and β_1, \dots, β_k in the cases of Poisson and Bernoulli firing times, possible differences in the vigor of animals searching for mates can be incorporated, cf. [13, p. 870]. However, it still remains to add other important features to the SEM model such as separation, births, deaths and offsprings in order to study the evolutionary aspects of mating behavior.

The key ideas. We would like to recapitulate that the following observations have been of fundamental importance throughout this paper:

- (i) In the case of definite mating upon encounter, $\mathcal{L}(\cdot)$ is measurable w.r.t. $\mathcal{L}(T)$ and $\tau = (\tau_s(\zeta))_{s \geq 1, \zeta \in Z}$ which are independent (see Theorem 2.3 and display (2.5)).
- (ii) In the Poisson and Bernoulli cases, we have the freedom to change P , $\alpha_1, \dots, \alpha_k$ and β_1, \dots, β_k as long as Π stays the same (see Remarks 3.2 and 4.2).

Indeed, the first one provides us with a convenient way of decomposing $\mathcal{L}(\cdot)$ into independent components; and the second one (which we have been referring to as the change-of-parameters technique) allows us to reduce the model to definite mating upon encounter if the corresponding fine balance condition is satisfied.

On generic EM laws. For Poisson firing times, the expectations $u_{ij}(t; x_1, \dots, x_k; y_1, \dots, y_k)$ and $u_{ij}^*(x_1, \dots, x_k; y_1, \dots, y_k)$, which were defined in (1.3) and (1.4), solve the recursive equations (3.3) and (3.4), respectively. When the EM law Π satisfies the Poisson fine balance condition, Theorem 3.6 provides formulas for (1.3) and (1.4). Similarly, in the Bernoulli case, (1.3) and (1.4) solve the recursive equations (4.3) and (4.4), respectively, and Theorem 4.6 provides formulas for them when Π satisfies the Bernoulli fine balance condition. One can also verify all of these statements by elementary combinatorics.

In either the Poisson or the Bernoulli case, if the corresponding fine balance condition is not satisfied, then we do not have formulas for (1.3) and (1.4), i.e., it is not known whether the aforementioned recursive equations can be solved. These constitute interesting but rather difficult open problems. Alternatively, one can try to formulate and prove categorical results regarding (1.3) and (1.4), such as Theorems 3.9 and 4.9 in which we use Π to characterize heterogamy/panmixia/homogamy for $k = 2$. It remains to figure out how these results can be generalized to $k \geq 3$.

Three interesting observations regarding mating preferences. First, as we have emphasized in Subsection 1.3, the terms panmixia, homogamy and heterogamy refer to the expected mating pattern and not to the mating preferences. Indeed, in the $k = 2$ case with Poisson or Bernoulli firing times, homogamy does not imply that (say) type-1 females prefer mating with type-1 males. For example, take

$$0 < p_{11} < p_{12} = p_{21} \leq \frac{1}{2} < p_{22} = 1$$

with $\alpha_1 = \alpha_2 = 0$ and $\beta_1 = \beta_2 = 1$. Then,

$$\pi_{11} + \pi_{22} > \pi_{12} + \pi_{21} \quad \text{and} \quad (1 - \pi_{11})(1 - \pi_{22}) < (1 - \pi_{12})(1 - \pi_{21}),$$

and we have homogamy both for Poisson and Bernoulli firing times by Theorems 3.9 and 4.9.

Second, in the $k = 2$ case, there exist mating preference matrices that give heterogamy for Poisson firing times and homogamy for Bernoulli firing times (or vice versa). For example, suppose

$$p_{12} = p_{21} = \frac{1}{2} < \frac{3}{4} \leq p_{22} < p_{11} + p_{22} < 1$$

with $\alpha_1 = \alpha_2 = 0$ and $\beta_1 = \beta_2 = 1$. Then,

$$\pi_{11} + \pi_{22} < \pi_{12} + \pi_{21} \quad \text{and} \quad (1 - \pi_{11})(1 - \pi_{22}) < (1 - \pi_{12})(1 - \pi_{21}),$$

and the claim follows from Theorems 3.9 and 4.9.

Third, let us assume that there exists $c \in (0, 1)$ such that $p_{ij} = c$ for every $i, j \in \{1, \dots, k\}$. In words, animals have uniform preferences over types. This case is rather similar to definite mating upon encounter, i.e., when $c = 1$. Therefore, one might predict that panmixia holds. This prediction is indeed true in the Poisson case with arbitrary $\alpha_1, \dots, \alpha_k$ and β_1, \dots, β_k since we get

$$\pi_{ij} = c(\alpha_i + \beta_j)$$

which satisfies the Poisson fine balance condition, and Corollary 3.7 applies. However, in the Bernoulli case with arbitrary $\alpha_1, \dots, \alpha_k$ and β_1, \dots, β_k , we get

$$\pi_{ij} = c(\alpha_i + \beta_j - \alpha_i \beta_j) = c[1 - (1 - \alpha_i)(1 - \beta_j)].$$

It is easy to check that this EM law satisfies the Bernoulli fine balance condition if and only if $(\alpha_i - \alpha_{i'})(\beta_j - \beta_{j'}) = 0$ for every $i, j, i', j' \in \{1, \dots, k\}$. Therefore, by Corollary 4.7, panmixia is equivalent to having $\alpha_1 = \alpha_2 = \dots = \alpha_k$ or $\beta_1 = \beta_2 = \dots = \beta_k$. Moreover, in the $k = 2$ case, it follows from Theorem 4.9 that there is homogamy when $(\alpha_1 - \alpha_2)(\beta_1 - \beta_2) < 0$ and heterogamy when $(\alpha_1 - \alpha_2)(\beta_1 - \beta_2) > 0$.

Asymptotics of the model. As we have mentioned in Subsection 1.3, Gimelfarb carries out his analysis of the individual and mass EM models in [13] after replacing all quantities such as $Q(T)$ with their expectations, and says that this approximation is justified by the law of large numbers (LLN) when the population is large. Therefore, it is natural to ask if the LLN and the central limit theorem indeed hold for the SEM model as x_1, \dots, x_k and y_1, \dots, y_k with (1.1) go to infinity in a suitable way. In a recent paper [14], we establish the functional versions of these limit theorems in the case of Poisson firing times, and describe the asymptotic mean (resp. the rescaled fluctuations) of $Q(\cdot)$ as the solution of an ODE (resp. SDE) system which we turn into Lotka-Volterra and replicator equations from population dynamics. We pay particular attention to the $k = 2$ case with $\pi_{12} = \pi_{21}$ and $x_1 = y_1$, and provide an explicit formula for the asymptotic (normalized) mating pattern without any fine balance assumption. We intend to do the same for Bernoulli firing times in the near future.

ACKNOWLEDGMENTS

We are indebted to A. Courtiol for introducing us to Gimelfarb's work on encounter-mating and for suggesting interesting problems. We also thank P. Diaconis and F. Rezakhanlou for valuable comments and discussions. O. Gün gratefully acknowledges support by DFG SPP Priority Programme 1590 "Probabilistic Structures in Evolution". A. Yilmaz is supported in part by European Union FP7 grant PCIG11-GA-2012-322078.

REFERENCES

- [1] S. Alpern and D. Reyniers. Strategic mating with homotypic preferences. *J. Theoret. Biol.*, 198(1):71–88, 1999.
- [2] S. Alpern and D. Reyniers. Strategic mating with common preferences. *J. Theoret. Biol.*, 237(4):337–354, 2005.
- [3] M. Andersson. *Sexual selection*, volume 27 of *Monographs in Behavior and Ecology*. Princeton University Press, second edition, 1994.
- [4] C. T. Bergstrom and L. A. Real. Towards a theory of mutual mate choice: Lessons from two-sided matching. *Evol. Ecol. Res.*, 2:493–508, 2000.
- [5] A. Courtiol, J. E. Pettay, M. Jokela, A. Rotkirch, and V. Lummaa. Natural and sexual selection in a monogamous historical human population. *Proceedings of the National Academy of Sciences*, 109(21):8044–8049, 2012.
- [6] C. S. Davis. *Statistical methods for the analysis of repeated measurements*. Springer Texts in Statistics. Springer-Verlag, New York, 2002.
- [7] K. Dietz and K. P. Hadeler. Epidemiological models for sexually transmitted diseases. *J. Math. Biol.*, 26(1):1–25, 1988.
- [8] A. M. Etheridge and S. Lemaire. Diffusion approximation of a multilocus model with assortative mating. *Electron. J. Probab.*, 16:no. 78, 2122–2181, 2011.
- [9] L. Etienne, F. Rousset, B. Godelle, and A. Courtiol. How choosy should I be? The relative searching time predicts evolution of choosiness under direct sexual selection. *Proceedings of the Royal Society B: Biological Sciences*, 281(1785), 2014.
- [10] W. J. Ewens. *Mathematical population genetics. I*, volume 27 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, second edition, 2004. Theoretical introduction.
- [11] J. H. Gillespie. *Population genetics. A Concise Guide*. The Johns Hopkins University Press, 1998.
- [12] A. Gimelfarb. Processes of pair formation leading to assortative mating in biological populations: dynamic interaction model. *Theoret. Population Biol.*, 34(1):1–23, 1988.
- [13] A. Gimelfarb. Processes of pair formation leading to assortative mating in biological populations: encounter-mating model. *American Naturalist*, 131(6):865–884, 1988.
- [14] O. Gün and A. Yilmaz. Fluid and diffusion limits for the Poisson encounter-mating model. Preprint, available at arXiv:1411.7220 [math.PR].
- [15] K. P. Hadeler. Pair formation. *J. Math. Biol.*, 64(4):613–645, 2012.
- [16] P. G. Higgs and B. Derrida. Genetic distance and species formation in evolving populations. *Journal of Molecular Evolution*, 35(5):454–465, 1992.
- [17] M. Kirkpatrick. Sexual selection and the evolution of female choice. *Evolution*, 36(1):1–12, 1982.
- [18] R. Lande. Models of speciation by sexual selection on polygenic traits. *Proc. Nat. Acad. Sci. U.S.A.*, 78(6, part 2):3721–3725, 1981.
- [19] C. J. Mode. Some multi-dimensional birth and death processes and their applications in population genetics. *Biometrics*, 18(4):543–567, 1962.

- [20] F. Mosteller. Association and estimation in contingency tables. *J. Amer. Statist. Assoc.*, 63:1–28, 1968.
- [21] A. K. Romney. Measuring endogamy. In *Explorations in Mathematical Anthropology*, ed. P. Kay, pages 191–213. Cambridge, 1971.
- [22] C. E. Taylor. Differences in mating propensities: some models for examining the genetic consequences. *Behavior Genetics*, 5(4):381–393, 1975.

ONUR GÜN, WEIERSTRASS INSTITUTE, MOHRENSTRASSE 39, 10117 BERLIN, GERMANY
E-mail address: `Onur.Guen@wias-berlin.de`

ATILLA YILMAZ, BOĞAZIÇI UNIVERSITY, DEPARTMENT OF MATHEMATICS, 34342 BEBEK, ISTANBUL, TURKEY
E-mail address: `atilla.yilmaz@boun.edu.tr`